THE EFFECT OF ACHIEVEMENT GOALS ON SELF-EXPLANATION AND TRANSFER: INVESTIGATING THE ROLE OF MOTIVATION ON LEARNING

by

Daniel M. Belenky

B.A., Syracuse University, 2005

Submitted to the Graduate Faculty of

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Daniel Matthew Belenky

It was defended on

November 5th, 2012

and approved by

Christian Schunn, Professor, Department of Psychology, University of Pittsburgh

John Levine, Professor, Department of Psychology, University of Pittsburgh

Vincent Aleven, Associate Professor, Human-Computer Interaction Institute, Carnegie Mellon University

Dissertation Advisor: Timothy Nokes-Malach, Assistant Professor, Department of Psychology, University of Pittsburgh

The Effect of Achievement Goals on Self-Explanation and Transfer: Investigating the Role of Motivation on Learning

Daniel M. Belenky, PhD

University of Pittsburgh, 2013

The role of student motivation in learning is an important issue for research to address, both for theoretical and practical purposes. The present research study tested a hypothesized behavioral pathway for previously documented benefits of student motivation – in particular, mastery-approach achievement goals – on learning and transfer (Belenky & Nokes-Malach, 2012). Achievement goals are the reasons students have for engaging in academic settings, such as wanting to develop their competence (mastery-approach goal), wanting to do better than their peers (performance-approach goal), or not wanting to do any worse than their peers (performance-avoidance goal). The present study manipulated achievement goals (mastery-approach, performance-approach, or performance-avoidance) that participants adopted while learning, and subsequently being tested on, basic statistical knowledge and procedures. It was predicted that mastery-approach goals would lead to higher levels of knowledge transfer. Additionally, talk aloud protocols were collected and coded to test hypothesis that mastery-approach goals would lead to more constructive learning processes, such as self-explanation. Finally, it was expected that the degree of self-explanations a student engaged in would be predictive of transfer. These hypotheses were not supported. Contrary to expectations, the performance-avoidance condition produced higher levels of transfer than the mastery-approach condition. Additionally, there were no differences between conditions in the amount of self-explanations generated. The amount of self-explanations was itself not predictive of transfer. These results are discussed in terms of possible ways the methodology may have reduced the

difficulty of transfer, as well as what the results may mean for achievement goal theory, more broadly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Research within cognitive psychology has led to many insights and theories about how people learn. These theories are sometimes adopted to help inform and improve educational practices (Bransford, Brown, & Cocking, 1999; Pashler et al., 2007). For example, cognitive psychology research on the testing effect (e.g., Roediger & Karpicke, 2006), spaced practice (e.g., Pashler, Cepeda, Rohrer, & Carpenter, 2007), analogical comparison (e.g., Gentner, Loewenstein, & Thompson, 2003), and the use of worked examples (e.g., Renkl, Stark, Gruber, & Mandl, 1998) have found ways to improve student learning, retention of knowledge, and transfer to novel situations or problems. Additionally, theories of skill acquisition have been used with great success to inform cognitive models underlying the design of intelligent tutoring systems (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995; Graesser, Chipman, Haynes, & Olney, 2005). The success of these models stems from their ability to accurately predict students' current state by decomposing the knowledge represented in a given problem into components, and tracking whether or not students possess those individual components and use them to correctly solve problems. Through this process, these tutoring systems can provide detailed feedback and guidance tailoring specific questions to the type of knowledge a student seems to be lacking. For instance, systems can be designed to prompt students to explain the principles that underlie some formulaic procedure (Aleven & Koedinger, 2002). Additionally, testing these models in real-life settings has allowed researchers to evaluate predictions and refine theories to be more applicable outside of the laboratory. However, even though cognitive psychology research has helped to develop many instructional principles, research conducted in classrooms still finds much variation in learning outcomes among students, even when using instructional practices based on these “best practices.”

To continue to develop theories of how students learn and improve educational outcomes, it is critical to incorporate aspects of student motivation. As any educator can tell you, how well a student learns from instruction can heavily depend on the student themselves; their attitudes, beliefs, goals, etc. In particular, these factors can influence the degree to which a student engages in particular learning behaviors. Cognitively grounded models of student learning, while powerful, do not typically examine the issue of what leads students to engage in a beneficial way or not. Incorporating motivation into these models is critical for developing more robust theories that can be applied to classroom settings, but to do so will require a thorough understanding of how motivation affects learning. The study described here directly investigates this issue. Specifically, the question being examined is how motivation influences learning behaviors, and what effect motivation has on what is learned. Ultimately, by drawing on the large bodies of empirical research on these matters to develop testable hypotheses, it will be possible to develop a model linking motivation and cognition to learning outcomes, which can guide future research questions, as well as inform educational practice. The present study begins to map out a possible path for such a research program.

In particular, the study reported here manipulates the motivational goals that students adopt, following research on achievement goal theory (Elliot, 2005), and examines what influence these goals have on student learning behaviors and outcomes, in terms of knowledge transfer. Achievement goal theory is currently the dominant theory within educational psychology for understanding student motivation in academic settings, and it posits that students may have a variety of goals in relations to establishing their competence; they may be focused on developing their competence (mastery-approach goal), or they may be more focused on demonstrating their competence (performance goal). While this theory has received a lot of empirical attention, few studies have directly observed how these goals lead to differences in learning behaviors and have generally relied on self-report methods or on experiments which only examine performance measures.

The current study extends prior research by testing the hypothesis that mastery-approach goals lead to increased levels of self-explanation. This is an important relationship to establish, as both mastery goals and self-explanation have been found to relate to better learning outcomes (Belenky & Nokes-Malach, 2012; Chi, Bassok, Lewis, Reimann, & Glaser, 1989). The current study tests a potential behavioral pathway from mastery-approach goals to self-explanation to increased transfer, linking together important theoretical contributions on how students learn from educational and cognitive psychology.

To begin, I will describe prior research on achievement goals and their relation to academic learning. Particular attention will be paid to the two strands of research that have been conducted, one correlational, which focuses on achievement goals at a course-based level, and one experimental, which manipulates achievement goals at a task-based level. Additionally, I will review relevant prior research which has begun trying to disentangle the effects of task-based and course-based goals. Finally, I will describe in detail the phenomenon of self-explanation, and argue why it may be expected that mastery-approach goals may lead to this type of learning behavior, before presenting a study which empirically investigates that relationship.

1.1 ACHIEVEMENT GOAL MOTIVATION

1.1.1 Achievement Goal Theory

Many theories of motivation exist, with different viewpoints for understanding what drives human behaviors. These range from those that focus on organisms maintaining homeostatic principles (Hull, 1943) to those that describe motivation in terms of the hedonic principle (i.e., to increase pleasure and avoid pain; Freud, 1920/1990), to becoming self-determined (Deci & Ryan, 1985), among others. One theory of motivation that has received a large amount of attention, particularly in regards to academic behaviors and performance, is achievement goal theory. This theory grew out of earlier theories of achievement motivation (McClelland, Atkinson, Clark & Lowell, 1953) and competence motivation (Harter, 1982).

At the heart of achievement goal theory is the construct of “achievement goals,” which are the self-reported reasons for how and why people engage in achievement situations (Elliot, 2005). An achievement situation is one that is based on competence in some skill or which relies on some body of specific knowledge relevant to that setting. As goals, they can be defined as “a cognitive representation of a future object that (an) organism is committed to approach or avoid” (Elliot & Fryer, 2008; pg. 244). The model of achievement goals that is currently predominant in the literature is a 2 (Mastery or Performance) x 2 (Approach or Avoidance) framework (Elliot & McGregor, 2001). Mastery goals concern the *development* of competence, while performance goals concern the *demonstration* of competence. Approach goals seek to achieve positive outcomes, while avoidance goals seek to avoid negative ones. The 2 x 2 nature of this framework results in four separable achievement goals; performance-approach, performance-avoidance, mastery-approach, and mastery-avoidance (see Table 1). Performance-approach goals deal with trying to demonstrate competence positively, which is usually operationalized as being in comparison to others, while performance-avoidance goals are focused on not performing worse than other. Mastery-avoidance goals focus on not missing opportunities to improve competence or not losing already existing competence, while mastery-approach goals center on developing competence.

Table 1

Achievement Goal Constructs and Representative Measures

	Goal construct	Example Questionnaire Item
Mastery-Approach	To develop and/or improve competence	*My aim is to completely master the material presented in this class*
Mastery-Avoidance	To avoid a loss of competence and/or to avoid missing an opportunity to improve competence	*My aim is to avoid learning less than I possibly could*
Performance-Approach	To demonstrate competence, particularly in relation to others	*I am striving to do well compared to other students*
Performance-Avoidance	To avoid demonstrating a lack of competence, particularly in relation to others	*My goal is to avoid performing poorly compared to others*

Note. Example questionnaire items from the Achievement Goal Questionnaire – Revised (Elliot & Murayama, 2008). Students are asked to rate the degree to which they agree or disagree with the statement on a Likert scale.

These goal orientations guide the interpretation of events in the achievement environment, and produce characteristic patterns of cognition, emotion, and behaviors (Kaplan & Maehr, 2007). Settings where these are active may be academic, such as in-class assignments or at-home test preparation (e.g., Church, Elliot, & Gable, 2001; Wolters, 2004), or non-academic, such as sports (e.g., Conroy, Elliot, & Hofer, 2003), or work environments (e.g., Schmidt & Ford, 2003; Wang, Chatzisarantis, Spray, & Biddle, 2002). This general model has been validated across a number of studies (e.g., Elliot & McGregor, 2001; Elliot & Murayama, 2008), widely used in various countries (e.g., Bong, 2009; Darnon & Butera, 2005; Murayama, Zhou, & Nesbit, 2009), across age groups (e.g., Bong, 2009; Urdan, 1997) and across contexts; for example, they have been measured within specific academic courses (e.g., Bong, 2004), for school, broadly (e.g., Finney, Pieper, & Barron, 2004), in athletic settings (see Ntoumanis & Biddle, 1999) and in the workplace (e.g., VandeWalle, Brown, Cron, & Slocum, 1999). Since these goals are considered orthogonal they can be studied independently (a "variable-centered" approach) or jointly (a "person-centered" approach; Tuominen-Soini, Salmela-Aro, & Niemvirta, 2011), although the majority of research follows a variable-centered approach, as will the present study. In a typical study which uses a variable-centered approach, all of the goals are measured via self-report questionnaires (see Table 1 for example items) and the independent contribution of each on an outcome measure is determined, using a range of statistical methods. These can range from simple bivariate correlations (e.g., Howell & Watson, 2007) to complex hierarchical linear models (e.g., Murayama & Elliot, 2009).

Mastery-approach, performance-approach, and performance-avoidance goals have been extensively studied, while mastery-avoidance goals are a relatively newer theoretical contribution (Elliot & McGregor, 2001), and have been the subject of fewer empirical studies. A consistent pattern of results regarding mastery-avoidance goals has yet to emerge, with some studies showing that they are related to positive outcomes (Madjar, Kaplan, & Weinstock, 2011; Richey & Nokes-Malach, 2013), and some demonstrating negative outcomes, relative to other goals (Van Yperen, Elliot, & Anseel, 2009). Given this lack of clarity on the expected contribution of mastery-avoidance goals, the work presented here will focus on the other three achievement goals that have received a more substantial amount of empirical and theoretical attention.

Performance-avoidance goals have been shown to consistently predict negative outcomes, in terms of behaviors, affect, strategies, and achievement. For example, higher degrees of performance-avoidance goals have been shown to relate to less organized study behavior and shallow processing (Elliot, McGregor, & Gable, 1999), test anxiety (Elliot & McGregor, 1999), lower interest (Harackewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008) and lower grades (Elliot & Church, 1997). While, more generally, seeking to avoid negative outcomes can be advantageous in certain environments (see research on prevention focus; Higgins, 1997), prior research and theory does not offer any indication that performance-avoidance goals may produce positive outcomes in learning environments. At best, someone adopting a strong performance-avoidance goal would be satisfied achieving a bare minimum level of competence. This may suffice in some fields of study, but, frequently, the goal of an educator is to produce relatively higher levels of competence, allowing for more complex problem-solving, reasoning, and transfer to a wider range of contexts.

Performance-approach goals have been related to both positive and negative academic outcomes. Positive outcomes linked with performance-approach goals include perseverance (Smith, Duda, Allen, & Hall, 2002), higher self-efficacy in some domains (e.g., writing; Pajares, Britner, & Valiante, 2000), use of both deep and surface learning strategies (Wolters, 2004), and achievement outcomes (see Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002, for a review). Negative outcomes observed include lower self-efficacy in some domains (Skaalvik, 1997), only surface learning strategies (Elliot & McGregor, 2001), and higher levels of extrinsic motivation (Smith et al., 2002). There also appear to be many interactions that influence whether performance-approach goals predict positive, negative, or no differences in student behaviors, affect, and achievement outcomes. For example, cultural differences (Liem, Lau, & Nie, 2008), evaluations of competence (Smith et al., 2002), and performance contingencies (Elliot, Shell, Henry & Maier, 2005) have been shown to moderate the effects of performance-approach goals on different outcomes.

Mastery-approach goals have shown the most promise for facilitating deeper conceptual learning and knowledge transfer. They have been associated with positive affective experiences, enjoyment, high self-efficacy, perceptions of competence, topic interest, better self-regulation and strategy use, and deeper processing of material (e.g., Ford, Smith, Weissbein, Gully, & Salas, 1998; Harackiewicz et al., 2008; Pintrich, 1999; Somuncuoglu & Yildirim, 1999). Additionally, a small body of research has documented a relationship between mastery-approach goals and knowledge transfer (Bereby-Meyer & Kaplan, 2005; Ford et al., 1998; Gist & Stevens, 1998; Kozlowski et al., 2001), although these have generally not been using academic domains (cf. Belenky & Nokes-Malach, 2012; in press).

1.1.2 Experimental and Non-Experimental Methods in Achievement Goal Research

The pattern of results discussed so far has been from studies using non-experimental methods. That is, in this line of research, achievement goals for a particular course are measured with an instrument such as the Achievement Goal Questionnaire – Revised (AGQ-R; Elliot & Muryama, 2008; see Table 1) and correlated to measures of antecedents of achievement goals (e.g., levels of interest, motives such as need for achievement, etc.) or outcomes, such as self-reported study strategies and achievement measures like grades (e.g., Somuncuoglu & Yildirim 1999). Sometimes, more advanced statistical methods are used to determine the relative contribution of achievement goals and certain mediating variables on particular outcomes (e.g., Grant & Dweck, 2003). These "course-level" achievement goals have been studied extensively, and a large volume of research has been published following this general method. For example, a recent meta-analysis of non-experimental achievement goal studies used a final sample of 243 studies, even after a rigorous selection process (Hulleman, Schrager, Bodman & Harackiewicz, 2010).

There is a second body of research on achievement goals that focuses on measuring the effect of task-based goals on a variety of performance and affective measures. Generally, this is done by manipulating the achievement goal a participant adopts for a specific task in a controlled experiment by framing the purpose of the experiment for the participant (e.g., Darnon, Butera, & Harackiewicz, 2007). Additionally, some studies state explicitly what a participant should aim to achieve during the experiment (e.g., Senko & Harackewicz, 2005). As a representative example, Elliot and Harackiewicz (1996) had participants work on puzzle tasks ("Nina puzzles"), and provided written information about the purpose of the study. Specifically, participants in the performance-approach condition read that "students are fairly comparable in their ability to solve Nina puzzles, but some students stand out because they do quite well on the puzzles. This session will give you the opportunity to demonstrate that you are a good problem solver." The performance-avoidance condition was given similar instructions, but which instead said "… but some students stand out because they do quite poorly on the puzzles. This session will give you the opportunity to demonstrate that you are not a poor problem solver." The mastery condition was instead focused solely on providing information on the students' competence, stating "When you have completed the four puzzles, you will be provided with information regarding the percentage of the total hidden Ninas that you found in today's session." Studies vary in the exact wording of manipulations, but they tend to all focus on framing the purpose of the experiment in such a way that students will adopt the target goal for the specific task.

However, little empirical or theoretical work has been done to disentangle the effects of course-based vs. task-based achievement goals, and they are frequently treated as interchangeable in the literature. This lack of clarity can be seen in how researchers conceptualize achievement goals. For example, Elliot (1999) wrote that "achievement goals… are used to directly regulate achievement behavior. Thus, achievement goals are construed as… the cognitive-dynamic representations that ***proximally*** influence achievement-relevant processes and outcomes" (pg. 174; emphasis added). In fact, the achievement goal construct was created partly in response to prior theories of achievement motives (McClelland et al., 1953), which were dispositional in nature (Elliot & Church, 1997). That is, they posited that individuals varied in their needs for achievement (approach motivation) and fear of failure (avoidance motivation), and this variation affected how people behaved in achievement settings. Achievement goals may be influenced by these underlying achievement motives, but are supposed to be more specific to the particular context (see Figure 1). The Achievement Goal Questionnaire (Elliot & McGregor, 2001) and its revised form, the Achievement Goal Questionnaire – Revised (Elliot & Murayama, 2008), ask students to consider the questions in light of the particular course they are taking, rather than measuring their global predispositions across all achievement settings. This achievement setting is thought to contribute to the adoption of achievements as an "environmental factor" (Elliot, 1999). However, in this line of research, "environmental factors" almost always refer to classroom or school-wide structures, rather than focusing on individual learning or performance contexts, such as a particular homework assignment, or a given textbook chapter.

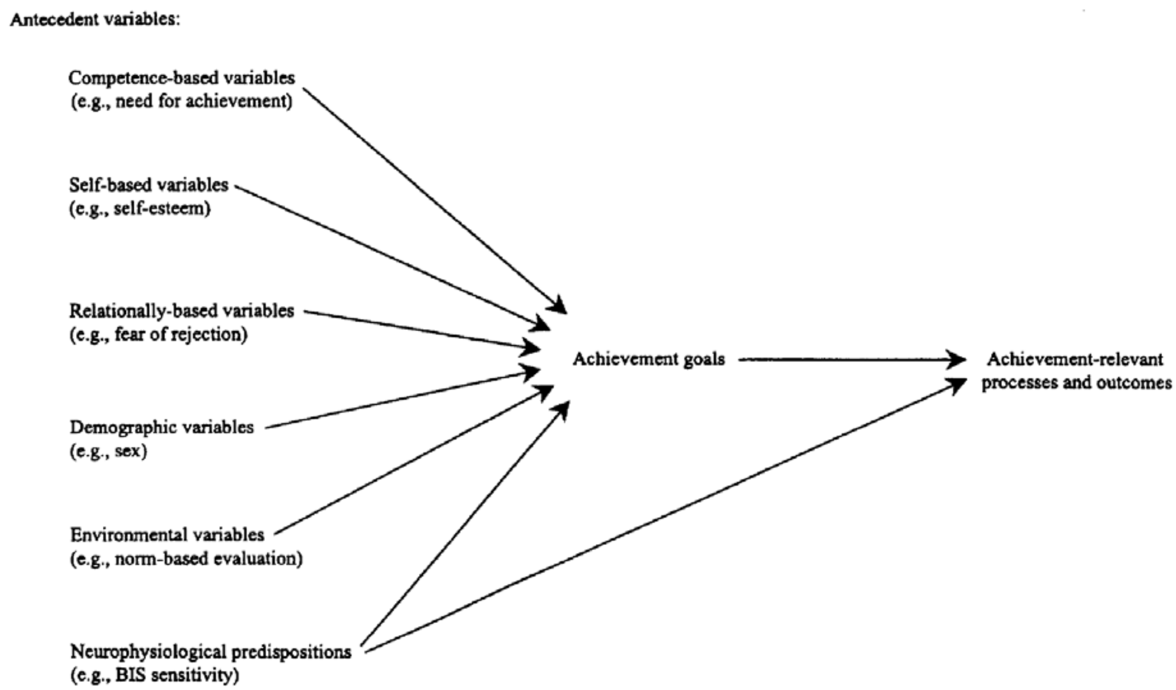

Figure 1. A model of achievement goals as proximal determinants of behavior. Adapted from "Approach and Avoidance Motivation and Achievement Goals," by A.J. Elliot, 1999, *Educational Psychologist, 34,* p. 177. Copyright 1999 by the American Psychological Association.

This lack of clarity brings into question the "proximal" nature of achievement goals. One can imagine that the distance from a course-level achievement goal to actual achievement-relevant behaviors is "further" than the distance from a task-based achievement goal to those behaviors. Research on achievement goals has not adequately addressed how a researcher is to know whether course-based or task-based goals are at play in any given learning scenario, and whether or not they produce similar effects. Although achievement goal theory posits that an achievement goal for a course and a task-based achievement goal would produce similar effects, this claim has not received much empirical support. More critically, achievement goals are thought to predict learning *behaviors*, something which has been examined using self-report measures in courses, but not using other methods to directly observe which goals predict which learning behaviors. Research in cognitive psychology has provided a great deal of evidence that the type of learning behavior a student engages in will determine the nature of the representation they form of the to-be-learned knowledge. As such, it is clear that this an important issue to address. Constructive learning behaviors, such as self-explanation and analogical comparison, have been shown to facilitate the formation of more abstract representations, which, in turn, help a student use that knowledge more flexibly when they have an opportunity to *transfer* their knowledge to a novel situation or problem. It is likely that achievement goals, when they have an effect on learning, act through the promotion of such learning behaviors (Elliot, 1999).

1.1.3 Prior Research on Experimentally-Manipulated Achievement Goals

To understand the effect of task-level goals, one can begin by examining the research that experimentally manipulates achievement goals. This research has produced some notable results, with studies indicating that manipulated mastery-approach goals are better for the development of intrinsic motivation (Elliot & Harackiewicz, 1996), that mastery-approach and performance-approach goals promote higher scores on IQ-like tests than performance-avoidance goals (Elliot, et al., 2005), and that mastery goals help students learn from disagreements with their peers, while performance goals help them learn when they are in agreement (Darnon, Butera, & Harackiewicz, 2007). Another example is that, in the face of failure, mastery-approach-oriented students view the difficulty as a challenge and increase their effort and the complexity of strategies employed (Elliott & Dweck, 1988).

However, reviews of these types of studies are less clear about the overall pattern of results. For example, Utman (1997) conducted a meta-analysis of 24 studies (analyzing 43 reported effect sizes) in which achievement goals were experimentally manipulated, and found a benefit for mastery goals compared to performance goals, but only when the tasks were complex.

Utman's meta-analysis is somewhat hampered, however, by the fact that this work did not take into account the more recent theoretical clarifications that have emerged, such as incorporating the approach/avoidance distinction between goals, or in clearer definitions of each goal. For example, the largest effect size (d = 2.66) Utman (1997) includes was contributed by a study reported in Amabile (1979). This famous study on the effect of expected evaluation on reducing creativity does not actually contrast mastery and performance goals. The methods state that the control conditions (which Utman used as the mastery condition) was told that the art they produced would not be evaluated, while that the evaluated conditions (which Utman used as the performance condition) were told that the art would be evaluated. While the expectation of evaluation is related to performance goal adoption (Ames, 1992), its absence is not enough to qualify as a mastery goal. Similarly, a very large effect size (d = 1.77) was contributed by the study reported in Benware and Deci (1984). This study manipulated whether participants were told that they should study with the expectation of being tested on the material (classified as performance by Utman), or whether they should study with the expectation that they would have to teach the material to another student (classified as mastery). While this study showed clear benefits for learning with the expectation of teaching the material, it does not seem directly relevant to the question of how achievement goal manipulations influence learning. Overall, the conclusion of this meta-analysis – that mastery goals are better than performance for complex tasks – needs to be tempered by the lack of theoretical precision used in constructing its sample.

Since Utman's meta-analysis was published, empirical and theoretical contributions have continued to be made, and some level of agreement has emerged as to what precisely constitutes an achievement goal, as opposed to another motivational construct (e.g., interest, challenge, naïve theory of intelligence, etc.). Although no clearly established "best practices" have been described, achievement goals tend to be manipulated in ways which correspond more closely to the definition of an achievement goal as the purpose one has for engaging in a competence-based ("achievement") setting. For example, mastery goals tend to be manipulated through a relatively consistent set of instructions. Specifically, these manipulations tend to have a central focus on the development of skill, albeit with some variability in exactly how this is introduced to participants. In some cases, it is focused on acquiring new knowledge, as in Darnon, Butera, and Harackiewicz (2007), which told the mastery condition that "You are here to acquire new knowledge that could be useful to you, to understand correctly the experiments and the ideas developed in the text, and to discover new concepts. In other words, you are here to learn" (pg. 64). In other cases, the focus is on the *process* of learning, such as in Graham and Golan (1991), which informed participants that "many people make mistakes on these puzzles in the beginning but get better as they go along… if you just concentrate on the task, try to see it as a challenge, and enjoy mastering it, you will probably get better as you go along" (pg. 189). This explicit mention of viewing the task as a challenge has become less common in more recent research, as researchers have begun to be more theoretically consistent by focusing on the development of competence as the core of a mastery-approach achievement goal. Additionally, the medium of the message varies, with some studies delivering the instructions verbally (e.g., Senko & Harackiewicz, 2005), and others providing text-based manipulations, included with other materials (e.g., Elliot et al. 2005).

Performance goals tend to be manipulated in one of two relatively consistent manners. One type of manipulation is to tell participants that their performance will demonstrate their ability in the given domain. This focus on revealing ability usually includes a statement that students tend to be either good (for performance-approach) or bad (for performance-avoidance) at the specific tasks being done. For example, Barker, McInerney, and Dowson (2002), told participants in the performance-approach condition that “people are either good at this activity compared to other kids their age or they are not. Your performance on this task will tell me something about how good you are at this kind of task.” While many studies used this sort of framing, it has clear limitations. For one, it can strongly interact with participants expectations and beliefs about the domain (although, this can also occur with the other method for manipulating performance goals, as described next). More critically, it conflates achievement goals with other motivational constructs; namely, entity theories of intelligence. The (frequently un-explicated) assumption is that informing someone that a task will reflect on underlying ability will prime an entity theory of intelligence in the person, which will lead them to adopt a performance goals (see Dweck & Leggett, 1988). However, this is an indirect path, and, while correlations have been observed between these constructs on course-based measures, it is not clear how a goal for given task will be influenced by adopting an entity theory, and whether that goal will necessarily be performance-oriented. In general, this “entity-based” technique for manipulating goals is more closely associated with an “ego-involving” achievement goal (Nicholls, 1984), which was a similar construct to what is currently considered a performance goal. However, the definition of an ego-involving goal differs slightly from the current conception of performance goals; performance goals are focused on the demonstration of competence relative to a normative standard, while ego-involving goals also included a larger focus on considering that competence as reflective of *ability* (Jagacinski & Nicholls, 1984). That being said, a body of research that has incorporated these aspects into manipulations has emerged, so it is clear that this focus on ability is efficacious in changing some aspect of student motivation, with concomitant changes in affect and behavior.

The current modal method for manipulating performance goals is to focus on a normative comparison, by telling participants that their performance on the task will be compared to others, so they should focus on trying to do better than others (in performance-approach manipulations) or on trying to not do any worse than other students (in performance-avoidance manipulations). One dimension that does vary is whether participants have a specific idea of what that goal entails – that is, if they are told what will be regarded as good or bad performance. For example, in Van Yperen (2003), participants were given a specific performance criterion which they should aim for, while in Elliot et al. (2005), they were told that the study would provide an opportunity "to demonstrate that you are an exceptional problem solver" (pg. 633). In the current study, a mixture of both of these methods (i.e., making salient for participants that the possibility of both good and bad performance exists, and that they should, in particular strive to either stand out for doing well, or to avoid standing out for doing poorly) are used to influence performance goals, as will be described in the methods.

The more recent research has stayed closer to these types of manipulations, which allow for more theoretical clarity. Linnenbrink-Garcia, Tyson, and Patall (2008) reviewed 27 articles (reporting 33 studies), most of which were more recent than those used by Utman (1997), and which were generally closer to matching the current conceptualization of mastery and performance goals. This review mainly focused on directly comparing mastery-approach and performance-approach goals, with a number of the studies reviewed not containing a control condition. In general, these studies failed to show a benefit for one goal over the other, with 70% showing null results. Among the remaining studies which did show an effect, about two-thirds found a benefit for mastery-approach goals, relative to performance-approach goals, with the other third finding the opposite pattern. The effects which were found tended to be large (i.e., greater than *d* = .8; Cohen, 1992), and some non-significant effects also had medium to large effect sizes, indicating a possible lack of statistical power in this type of research. The results of the experimental manipulations of achievement goals reviewed in Linnenbrink-Garcia, Tyson, and Patall (2008), as well as those reported in other published studies, find a relatively consistent pattern of mixed results. That is, a large number of studies have results that do not cleanly differentiate between performance-approach and mastery-approach goals. Studies that include manipulations of performance-avoidance goals do tend to find that, relative to the two approach goals, task performance is lowered.

However, these studies did not examine whether the different goals lead to different learning *behaviors*. Different task-based achievement goals, which are supposed to be proximal influences of behaviors in achievement settings, should lead to different learning strategies, which should, in turn, lead to the construction of different types of knowledge representations. In particular, it is possible that the pattern of results reflects the fact that both mastery-approach and performance-approach lead to beneficial types of learning behaviors (e.g., persistence, attention, etc.), but that only mastery-approach goals promote deeper learning processes, such as self-explanation, or analogical comparison. Prior studies may not have been using performance measures that were sufficiently sensitive to detect these sorts of different types of knowledge representations, as evidenced by the relatively few published works that looked at the effect of achievement goals on transfer. The next section will discuss in detail one such prior study, which will be used as a model for the current research.

1.1.4 Prior Research on Mastery-Approach Goals and Transfer

A recent study investigated how course-level and task-level achievement goals interact with different forms of instruction to influence knowledge transfer (Belenky & Nokes-Malach, 2012). In this work, course-level achievement goals were measured before students performed learning activities about basic statistical concepts in one of two randomly assigned conditions. In one condition (called "Tell-and-Practice"), students were given direct instruction, in the form of worked examples, on procedures they could use during the learning activities (i.e., a mean deviation formula and a graphical method for standardizing scores). In the other condition (called "Invention"), students were prompted to *invent* procedures to solve the same problems. These invention activities have been shown to engage students in more effortful, constructive cognitive processes (Roll, Holmes, Day, & Bonn, 2012), and to facilitate learning in such a way that benefits transfer (Schwartz & Martin, 2004; Kapur, 2010). In particular, Schwartz and Martin (2004) found that invention students were more likely to successfully transfer in a "Transfer as Preparation for Future Learning" paradigm. This experimental paradigm is based on a theoretical conception of knowledge transfer which conceptualizes the ability to transfer knowledge as comprised of both transferring knowledge *in* to new situations where it can help learn the new material better, as well as transferring that new knowledge *out* to novel problem solving scenarios. This can be contrasted with some prior conceptions of transfer, which focused exclusively on the ability to transfer knowledge *out*, directly from instruction, to help solve a new problem.

To measure this "transfer as preparation for future learning," a double-transfer paradigm was used (Schwartz & Martin, 2004; see Figure 2). This paradigm randomly assigns participants to an instructional condition (Tell-and-Practice or Invention, for example), and then within each of those conditions, randomly assigns half of the participants to receive an embedded learning resource. This “extra” learning resource pertains to the critical transfer problem, and includes information that helps to solve that problem. For example, in studies that teach basic statistics, the worked example may demonstrate the formula for computing a standardized score, while the transfer problem asks participants to decide which of two scores from different distributions is more impressive. Transfer occurs if a participant uses the knowledge from the embedded resource on the transfer problem. If there are differences between the instructional conditions on their ability to transfer from the embedded resource, this indicates that the different types of instruction produced different types of learning in the first place. That is, the ability to transfer out is related to how well the students transferred in to the embedded resource.

Thus, in this paradigm, transfer is conceptualized as the ability to solve a new problem after an opportunity to learn the relevant concept and related procedure. Barnett and Ceci (2002) classify the “distance” of transfer in terms of what content is transferred, and how different the context of transfer is from the original learning. To use their terms, the content that transfers in this paradigm is a “principle or heuristic” (Barnett & Ceci, 2002, pg. 621); specifically, successful performance requires learning the approach of calculating a standardized score from the worked example and transferring it to the critical problem. While it may appear to require the transfer of a problem-solving procedure only, which would be considered somewhat “nearer” transfer, the many (incorrect) intuitive approaches that students use on the pre-test and post-test provides evidence that participants lack more than just a formulaic procedure. Solving the problem without using some approximation of a standardized score indicates that the problem has been incorrectly represented, not just that the wrong formula was applied. As such, this assessment should be considered as an instance of conceptual transfer, as only by transferring the concept of "standardization" can a student successfully solve this problem.

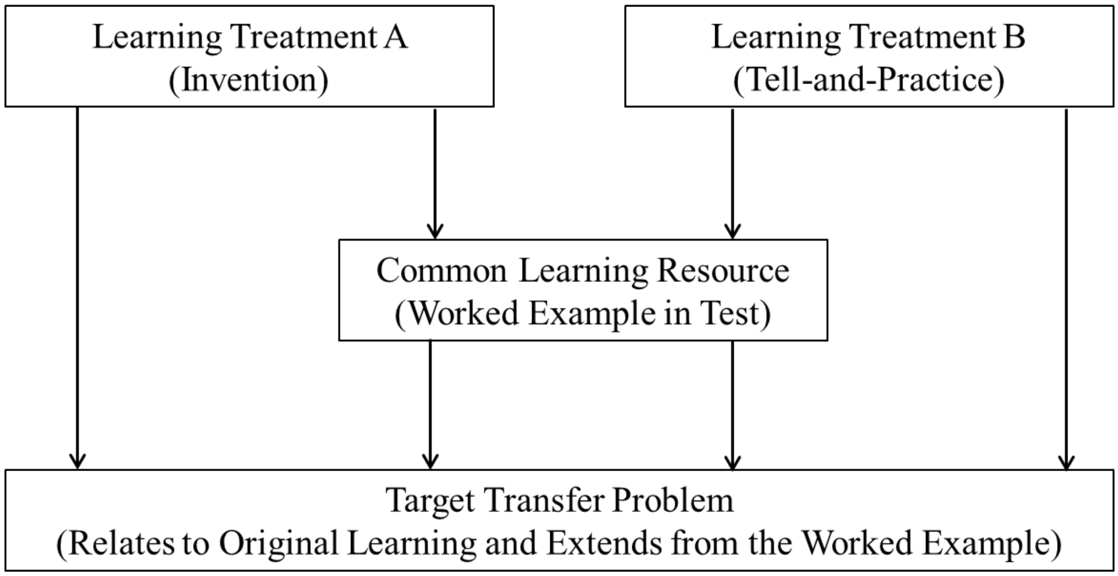

Figure 2. The "double-transfer" paradigm used in Belenky and Nokes-Malach (2012), which is based on Schwartz and Martin's (2004) methods.

Belenky and Nokes-Malach (2012), following the prior research method closely (Schwartz & Martin, 2004), found interesting results relating transfer to both course-level and task-level achievement goal motivation. Specifically, the higher a student's self-reported course-level mastery-approach goals for math were, the more likely she was to transfer. However, this effect interacted with the type of instruction a student received (see Figure 3). For students in the tell-and-practice condition, this positive, linear relationship between course-level mastery-approach goals and likelihood of transfer was strong, such that every unit change in mastery-approach goals corresponded to the odds of successfully transferring increasing by a factor of 4.8. However, for the invention condition, the relationship was greatly moderated. In this condition, every unit change in mastery-approach goals corresponded to an increase in the odds of transfer by a factor of only 1.09. By examining the likelihood of transfer for a hypothetical student low in course-level mastery-approach goals across conditions, this difference can be understood quite plainly; in the tell-and-practice condition, this student would be very unlikely to transfer, while, in the invention condition, they would be about as likely to transfer as a student high in course-level mastery-approach goals.

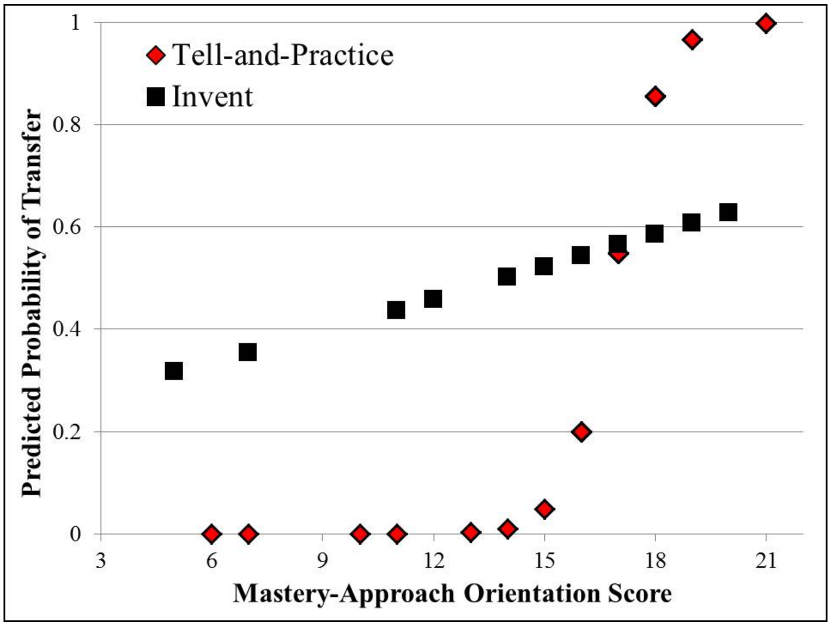

Figure 3. Transfer as a function of course-based mastery-approach goals for invention and tell-and-practice activities.

The reason for this pattern of results may have to do with the goals students adopted as a result of the instructional activities. Task-based mastery-approach achievement goals were assessed using a short questionnaire, administered directly after the learning activities. Students who completed invention activities reported higher levels of task-based mastery-approach achievement goals (d = .79). Although no statistical relationship between the degree to which a student endorsed task-based mastery-approach goals and transfer was observed, it is possible that the higher degree of task-based mastery-goal adoption is at least partially responsible for tempering the effect of course-based goals on transfer.

Stated another way, Elliot (1999) has said "if the achievement setting is 'strong' enough…, it alone can establish situation-specific concerns that lead to goal preferences for the individual, either in the absence of a priori propensities or by overwhelming such propensities." In this work, it is possible that proximal mastery-approach goals, which influenced the type of learning behaviors and, ultimately, the likelihood of transfer, were available to students from two different sources; their existing course-based goals, and task-based goals adopted as they completed the learning activities. In the tell-and-practice condition, the structure of the task was straightforward, and, if students did not have any desire to engage with the materials in a way that increased their competence, they were not required to. However, this was not an option for the invention students, who, by necessity, had to focus on increasing their competence simply to attempt to solve the problem.

To summarize this prior research, mastery-approach goals – whether task-based or course-based – seemed to increase the likelihood of transfer. In particular, higher levels of course-level mastery-approach goals for math increased the likelihood of transfer. However, for those low in course-level mastery-approach goals, a stark difference between invention and tell-and-practice was observed, such that those low in course-level mastery-approach goals were much more likely to transfer in the invention condition than in the tell-and-practice. This result, considered in light of the finding that invention led to higher levels of task-based mastery-approach goals, lends support to the idea that adopting a mastery-approach goal is beneficial for learning, whether that goal comes from existing course-level goals or emerges out of interaction with a particular type of learning activity. The few other research studies that have examined the link between achievement goals (whether task-based or course-level) and transfer have found similar patterns of results (see Pugh & Bergin, 2006). That is, mastery-approach goals, whether experimentally-induced for the particular task (Gist & Stevens, 1998), assessed as more general orientations (Ford, Weissbein, Smith, Gully, & Salas, 1998), or both (Bereby-Meyer & Kaplan, 2005; Bell & Kozlowski., 2002), are related to knowledge transfer from training to some other problem or setting. However, aside from the Belenky and Nokes-Malach studies (2012; in press), no other studies have examined this issue using academically-relevant materials and course-based achievement goals.

In particular, the empirical research on achievement goals has tended to focus on performance (sometimes using transfer as one measure of performance) rather than on *learning*, per se. As a result, research has not directly investigated what influence task-level goals have on learning behaviors. For example, Belenky and Nokes-Malach (2012) showed that task-based mastery-approach goals may be contributing to transfer, but did not analyze by what mechanism this benefit might be realized. To address this question, the current research adopts the standard research practice from cognitive psychology of gathering talk-aloud protocols to analyze for evidence of constructive processing, such as self-explanation. The following section will examine research on self-explanation, and the ways in which this paradigm can be used to address the issue of how achievement goals influence learning.

1.2 SELF-EXPLANATION

An important aspect of successful student learning, particularly in academic domains such as algebra and statistics, is understanding the justification for why procedures you are using work. That is, when a student learns successfully, she not only acquires proficiency in the to-be-learned skill, but, in addition develops an understanding for how that procedure works, as well as why and when it is appropriate to use that procedure. When a student attempts to develop this type of understanding, she can be said to be engaging in “sense-making” (Koedinger, Corbett, & Perfetti, 2012). It is possible to structure learning activities to try and promote sense-making (e.g., Ainsworth & Loizou, 2003; Chi et al., 1989), but it is also a type of cognitive activity that a student can spontaneously engage in on their own. Research has demonstrated that such “self-explanations” can benefit student learning, by clarifying when particular procedures apply (and when they do not), what effect each step of the procedure has, how the procedure helps achieve the goals of the problem, and how all of these goals and procedures relate to more general principles. This process is particularly important when students learn from worked examples (Chi, DeLeeuw, Chin, & Lavancher, 1994; Nokes-Malach, VanLehn, Belenky, Lichtenstein, & Cox, in press; Renkl, 1997).

Perhaps the most well-known example of research on self-explanation is the study reported in Chi et al. (1989). In this study, ten university students learned about physics concepts by reading a commonly used introductory physics textbook. As is typical, the chapters contained a mix of declarative text and worked examples that demonstrate problem solving procedures. In one critical chapter, the experimenters ensured that students had mastered the declarative portions of the chapter (through answering declarative and qualitative questions correctly) before letting the students study the worked examples. When students studied the worked examples, they were asked to talk-aloud, saying whatever they were thinking, and were prompted to “overtly explain to themselves what they understand, after reading every line of a worked example” (pg. 151). This talk-aloud methodology has been used widely in cognitive psychology, and is a powerful tool for investigating procedures students use when solving problems. In particular, the researchers focused their analyses on the explanations students generated for the various problem-solving steps. As Chi et al. (1989) state, “We postulate that explanations can reveal students’ understanding by showing whether or not they know: 1) The conditions of application of the actions; 2) The consequences of the actions; 3) The relationship of actions to goals; 4) The relationship of goals and actions to natural laws and other principles” (pg. 151).

Analysis of these talk-aloud protocols revealed stark differences between “Good” and “Poor” students, as defined post-hoc based on their performance on post-test measures. Good students spent more time and produced more lines of speech while studying examples, producing a larger amount of unique thoughts on the topics (“idea statements”). More of these idea statements were self-explanations, classified as a statement that served to demonstrate one of the four features discussed in the previous paragraph. Additionally, as good students monitored their own comprehension, they made statements of understanding and failure to comprehend in equal proportions, while poor students made very few statements indicating they did not understand. Finally, good students re-read very few lines from prior learning materials, seeking targeted information to address a specific issue or point of confusion. In contrast, when poor students re-read learning materials, they tended to read many lines, indicating that they were searching for something that would help them to find a solution.

In sum, self-explaining is a very beneficial learning behavior for a student to engage in. Many experiments conducted in the past 20 years have found that a variety of instructional techniques can be used to prompt students to self-explain, and that doing so leads to greater learning gains than asking students to re-read text, or to summarize it (e.g., Aleven & Koedinger, 2002; Chi et al., 1994; Renkl et al., 1998). In particular, these techniques help students develop a deeper understanding, promoting knowledge transfer. While some research has examined the teaching of self-explanation as a skill and a learning strategy (e.g., Wong, Lawson, & Keeves, 2002), the body of research which has prompted students to self-explain conceptually difficult material and found benefits for learning suggests that whether a student “possesses” the skill of self-explaining is likely not the critical issue. Rather, what is most important is whether the student *chooses* to use it or not. In the absence of prompting, what leads some students to engage in self-explanation while others do not? The current study investigates the hypothesis that self-explanation behaviors are partly determined by student motivation, and, in particular, by mastery-approach goals. This possibility has been hinted at before, but without the theoretical clarity that a framework, such as achievement goal theory, provides. For example, Chi et al. (1989) state “…we view the greater amount of protocols produced by the Good students as a natural consequence of *wanting to understand* the solution example better...” (pg. 160, emphasis added). The possible link between self-explanation and mastery-approach goals will be addressed next.

1.2.1 Self-Explanation and Mastery-Approach Goals

While the prior studies that found a relationship between course-level achievement goals and transfer did not collect talk-aloud protocols, they did show a relationship to “deep” learning outcomes, such as knowledge transfer. Additionally, prior correlational research has shown that self-reported course-level mastery-approach goals are related to deeper learning strategies, as measured by instruments such as the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich & DeGroot, 1990), which ask students to rate their own study behavior by responding to Likert-scale items. Questions that assess deep learning strategies tap concepts that are closely related to the idea of self-explanation (i.e., generating for oneself functional procedural knowledge from declarative knowledge). For example, these deep learning strategies encompass processes such as elaboration, critical thinking, and metacognitive self-regulation (see Table 2).

Table 2

Example Measures for Learning Strategies Related to Mastery-Approach Goals

	Sample Items	
Elaboration	*I try to understand the material in this class by making connections between the readings and the concepts from the lectures.*	*When reading for this class, I try to relate the material to what I already know.*
Critical Thinking	*When reading for this class, I try to relate the material to what I already know.*	*I often find myself questioning things I hear or read in this course to decide if I find them convincing.*
Self-Regulation	*I ask myself questions to make sure I understand the material I have been studying in this class.*	*I try to change the way I study in order to fit the course requirements and the instructor's teaching style.*

Note. Items from the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich & DeGroot, 1990). Students are asked to rate the degree to which they agree or disagree with the statement on a Likert scale.

These sorts of processes can be observed using talk-aloud protocol analysis, as described earlier. In particular, one can code for certain types of cognitive behaviors and procedures that may indicate a deeper level of processing. For example, deeper processing is evident when students generate elaborations which reflect inferences that provide underlying reasons or justifications for problem-solving steps, or bring attention to unresolved questions, or draw comparisons across learning opportunities to reflect a level of processing. Higher degrees of these types of statements would reflect more effort towards understanding the material deeply that, presumably, would lead to better learning outcomes.

It seems likely that in cases where mastery-approach goals increase "deep" learning outcomes, they do so at least partly through the mechanism of self-explanation. As quoted earlier, it may be considered "a natural consequence of wanting to understand" (Chi et al., 1989, pg. 160). As such, I predict that task-based mastery-approach goals will lead to more self-explanation, which in turn should lead to better learning outcomes, such as more transfer.

1.3 HYPOTHESES

The present study examined the effect of manipulated achievement goals on goal adoption and knowledge transfer from instruction. Talk-aloud protocols were collected as participants worked on the learning and test materials, so analyses are conducted on the types of cognitive strategies and processes participants used. The following research questions are addressed, with specific hypotheses offered for each of them.

1.3.1 Mastery-Approach Goals and Transfer

The central question of this study is as follows; do task-based manipulations of mastery-approach goals improve learning, as measured by transfer? Given past research for the benefits of mastery goals for transfer (e.g., Bereby-Meyer & Kaplan, 2005; Ford et al., 1998) and, in particular, past research with the same type of learning and test materials (Belenky & Nokes-Malach, 2012; in press), it seems likely that a task-based mastery-approach goal would lead to better transfer outcomes than other achievement goals. The specific hypothesis being tested was as follows:

H1: Task-based mastery-approach goals will lead to higher levels of transfer than task-based performance-approach goals, which will produce higher levels of transfer than task-based performance-avoidance goals.

Given that, in prior research, students high in course-level mastery-approach goals were more likely to transfer, task-based mastery-approach goals were also expected to facilitate transfer. Additionally, the relatively few studies that have examined the link between task-based mastery goals and knowledge transfer have found a positive relationship (see Pugh & Bergin, 2006). Though those studies did not focus on the learning of standard academic materials, the learning process related to mastery-approach goals should produce similar effects of knowledge transfer. Prior research also suggests that, while performance-approach goals produce mixed patterns of results, these goals tend to promote more positive learning behaviors and outcomes than do performance-avoidance goals. This leads to the prediction that more success transferring will be observed for the performance-approach condition, compared to performance-avoidance.

1.3.2 Mastery-Approach and Self-Explanation

This research will also address a second research question, “Do mastery-approach goals lead to increased self-explanation?” To assess this question, talk-aloud protocols were collected and analyzed for evidence of constructive cognitive processes. The specific hypothesis being tested was:

H2: Mastery-approach achievement goals will lead to more self-explanations than performance-approach goals, which will produce more self-explanations than performance-avoidance goals.

This prediction is supported by earlier theory and research. Specifically, research on course-level achievement goals has found that mastery-approach goals are related to self-reported use of more constructive cognitive strategies, like self-explanation. Additionally, as quoted from Chi et al. (1989) earlier, individual differences in the degree to which students self-explain are likely influenced by a student's desire to understand the material. Performance-approach goals are hypothesized to produce more self-explanations than performance-avoidance due to the patterns observed in the research literature. Specifically, performance-approach goals have produced somewhat mixed results, showing that this goal is sometimes related to deep learning strategies, and other times shallow, as well as sometimes being related to better performance, and other times related to worse performance. Performance-avoidance goals, in contrast, are consistently found to promote shallow learning strategies and poor performance outcomes.

1.3.3 Self-Explanations and Transfer

Prior research indicates that self-explanation helps transfer. As such, I expect that, regardless of condition, students who self-explain more will perform better on transfer measures.

> *H3: Higher levels of self-explanations and other constructive elaborations will predict transfer.*

It is also possible that self-explanations can serve different purposes for different achievement goals. Specifically, following the "learning agenda" hypothesis (Senko & Miles, 2008), it is possible that all students may self-explain, with important differences being related to the *content* and *location* of the self-explanations. That is, according to the learning agenda hypothesis, students with a mastery-approach goal tend to focus on the aspects of the learning material they find interesting, while performance-approach goals lead students to focus on learning aspects of the materials that they believe will help them be evaluated well. In a normal classroom, this difference may manifest itself in different modes of studying. A mastery-approach oriented student might be likely to study those aspects of the materials they find interesting in depth, perhaps seeking out additional readings on those topics. A performance-approach oriented student would be likely to consider what aspects of the material are likely to be assessed, and focus on learning those well enough to score highly on the test. In the current study, this may be expressed in different patterns of explanations for specific aspects of the materials. That is, it is possible that certain aspects of the learning materials which will, ultimately, aid in transfer, may correctly be picked out as the most relevant features and have extra attention devoted to them in the performance-approach condition.

2.0 THE PRESENT EXPERIMENT

The current study follows the methods of prior research (Belenky & Nokes-Malach, 2012; Schwartz & Martin, 2004) to explore the effect of task-based achievement goals on learning behaviors and knowledge transfer. To do so, a manipulation of achievement goals was utilized, and participants were recorded talking aloud as they worked on a set of learning and test materials about basic statistical procedures. This chapter will describe the research method of the present study in detail (see Figure 4 for an overview of the design, materials, and procedure).

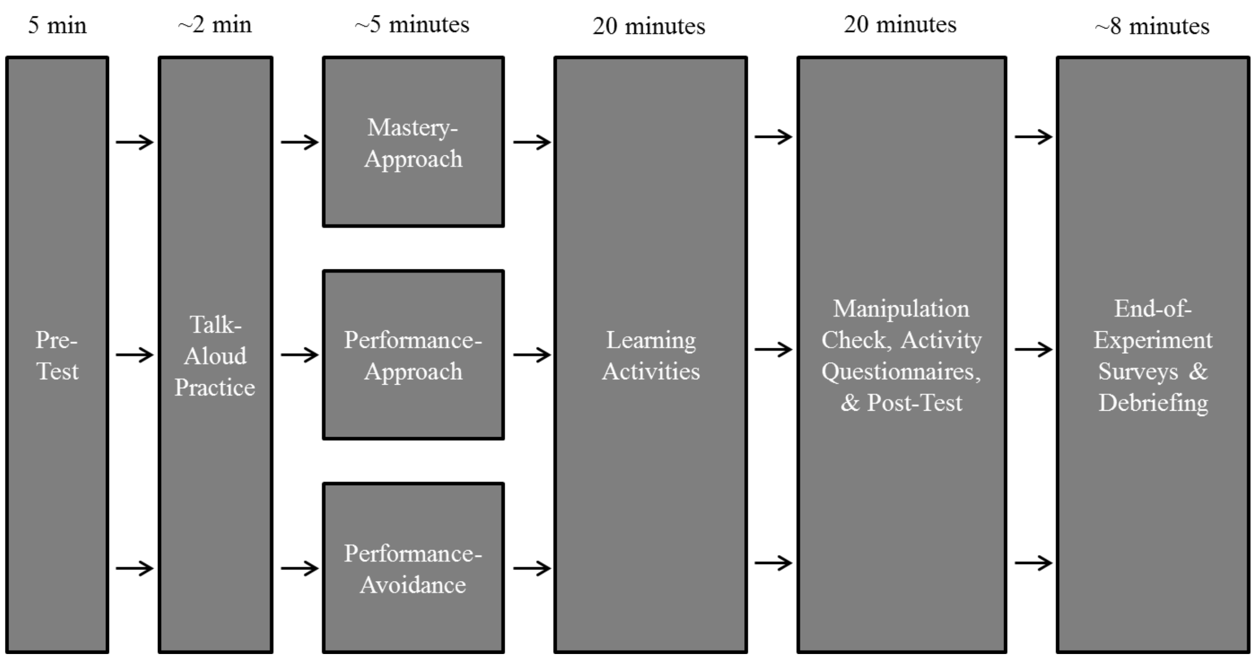

Figure 4. Representation of the study procedure.

2.1 METHOD

2.1.1 Design

This study is a between-subjects experiment with three conditions: mastery-approach, performance-approach, and performance-avoidance. All participants completed the same learning activities, test questions, and questionnaires, with the only difference between conditions being the framing given for the experiment.

2.1.2 Participants

One hundred and fifteen University of Pittsburgh undergraduates participated for course credit. Ten of these participants are not included in the analyses. Eight were non-native English speakers (one of whom did not complete the study because of language difficulties). These participants were removed because their talk-aloud protocols could not be analyzed. One participant did not want to continue after being given the task-framing manipulation, while another guessed the purpose of the experiment after the task-framing manipulation; therefore, data was not collected from either of these participants. After removing these, each condition had 35 participants. Of these 105 students in the final sample, 69% were freshmen, 20% were sophomores, 7% were juniors, 4% were seniors, and only 1 student reported as something other than those. Fifty percent were females and 50% were males. One of these students (in the mastery-approach condition) did not have their talk aloud data recorded due to experimenter error; this participant's data is included in analyses of learning, test performance, and motivation measures.

2.1.3 Materials

The materials for this study were adapted from prior research on transfer as preparation for future learning and motivation (Belenky & Nokes-Malach, 2012; Schwartz & Martin, 2004). These consisted of a pre-test, learning activities, a post-test, and motivation measures, described next. All of these materials were delivered in binders that were waiting on the desk when the participant entered the room in which the experiment was conducted.

2.1.3.1 Pre-Test

Test materials (both for the pre-test and post-test) were also adapted from prior research. The pre-test consisted of a procedural fluency measure, a transfer problem, and a graphical representation problem. The procedural fluency measure provides participants with a small data set (eight numbers) consisting of small single- and double-digit numbers, and asks the participant to calculate the mean, median, mode, and mean deviation. This was done to evaluate participants' prior knowledge of these terms and procedures.

Following this was a transfer problem, which dealt with the concept of using standardized scores to compare values from two different distributions. Specifically, the problem text was:

> *Joe Smith and Mike Brown were baseball players from different eras (Smith played in the 30s, Brown in the 60s), but they were both power hitters famous for their ability to hit homeruns. Two friends were arguing whether Joe Smith or Mike Brown really had more power. Joe Smith's longest homerun was 540 ft. That year, the mean homerun among all players was 420-ft long, and the average deviation was 70 ft. The average deviation indicates how close all the homeruns were to the average. Mike Brown's longest homerun was 590 ft. That year, the mean homerun was 450 ft, and the average deviation was 90 ft. Who do you think showed more power for his biggest homerun, Joe Smith or Mike Brown? Use math to help back up your opinion.*

Correctly solving this problem requires calculating a standardized score, so that the degree to which each homerun is exceptional can be determined, given the distribution it comes from. This problem is constructed in such a way that students may choose solutions that reflect particular misconceptions. Specifically, students may rely on intuitive reasoning and say that one homerun is longer than another so that one is “objectively” better. There are many reasons why such an approach is incorrect, but it is certainly true that standards change over time and relying solely on absolute values will often produce suboptimal solutions. Students who attempt to follow instructions and “use math to help back up your opinion” sometimes rely on what may be called “distance from deviation” reasoning. Students using this particular approach will add the deviation to the mean for a given distribution, and the compare the difference between the exceptional score and this “mean plus deviation” value. This approach leads students to say that both Mike Brown and Joe Smith showed equal power, as both are 50 feet longer than their “mean plus deviation” value. Only responses that reflected the calculation of a standardized score, whether expressed graphically or numerically, were counted as correct, and received a score of two. Problems which demonstrated reasoning that reflected conceptually accurate approaches, but did not use standardized scores, were given a score of one (e.g., “Joe showed more power because Mike’s year had scores all over the place, but Joe hit it further when most people bunched around a lower number). A score of zero was assigned for all other responses. Two raters (the author and a research assistant) coded the entire data set (κ =.76; the author’s codes were used in analyses)..

Following this problem, participants received a graphical representation problem. This problem gave students two different data sets, representing the win totals for two different high school football teams, and asked them to draw a graph for each, so that they can make a decision about which of the two teams is better. This problem was included mainly because pilot testing revealed that participants would not be able to finish all three of these problems in the five minutes allotted. There is evidence that achievement goals are more influential in situations where success or failure is uncertain for participants, and prior research has integrated the idea of students not finishing directly into instructions (e.g., Senko & Harackiewicz, 2005). Thus by giving them more material than they can reasonably finish in the time allotted was used as a way to ensure that participants' would feel some level of uncertainty regarding their expectations for success in the rest of the experiment.

2.1.3.2 Learning Materials

The learning materials were adapted from the "Tell-and-Practice" materials used in prior research (Belenky & Nokes-Malach, 2012; Schwartz & Martin, 2004). Tell-and-practice was used because prior research had shown that course-based mastery-approach goals had more influence in this condition than in the invention conditions, as discussed in section 1.1.4. These materials are comprised of worked examples and problems that introduce, model, and provide opportunities for practice on two basic statistical concepts; mean deviation and standardized scores.

Here you will learn a technique for calculating how reliable a set of data is, by computing how variable it is. To do this, you must first calculate the mean. For example, take the data set:

x
26
27
27
28
29
36
45
47
47
48

The mean of this set is 36, which is calculated by adding up all of the elements and dividing by the number of data points in the set. That is, we add 26 +27 + 27 + 28 + ... + 48, which equals 360. We then divide by 10, since there are 10 values in the data set. This gives us a mean of 36.

x	|Mean- x|
26	10
27	9
27	9
28	8
29	7
36	0
45	9
47	11
47	11
48	12

To calculate the variability, you need to calculate how far each variable is from the mean. Because we do not want to end up with negative values, you should take only the absolute value of the difference of the data point from the mean. We can call this value the data point's "deviation." Doing so gives us the column of data shown to the right.

Now, you can find the mean of deviations. This will give you the average deviation, a measure of how variable the data is. In this case, the sum of all of the deviations (10 + 9 + 9 +... + 12) equals 86. We divide this by 10 (since there are 10 values), and find that this data set has an average deviation of 8.6.

Figure 5. Worked example for calculating mean deviation**.**

Participants first received a worked example on how to calculate mean deviation (see Figure 5). This worked example demonstrates how to calculate a mean deviation value for a data set. This is followed by a learning activity problem ("Pitching Machines"). This problem presents data from four pitching machines and asks the participant to decide which of the four is the most reliable (see Figure 6). The datasets for these four pitching machines are designed in such a way that contrasting between them should help focus participants' attention to the critical features of the mean deviation formula (and their conceptual underpinnings), such as the number of data points, the spread, the direction of the misses, etc. However, with the tell-and-practice nature of the activity, these aspects could be ignored in favor of a "plug-and-chug" method.

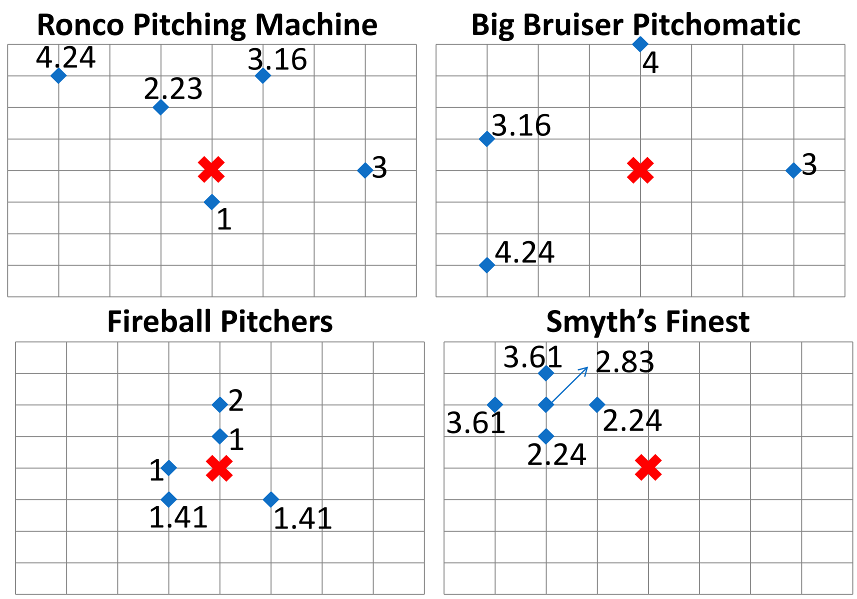

Figure 6. The "Pitching Machines" problem, which asked students to decide which of the machines was most reliable.

After completing this problem, participants moved on to the next section. Here, they read a two-page worked example that demonstrates how to compare two scores from different distributions to one another. Specifically, the problem was about two students in different classes who want to know who did better on a test, given that their teachers may grade differently. Over the two pages, the worked example shows the participant how to draw a histogram for each of the classes, and then how to map the given information about means, mean deviations, and the particular students' scores onto the histogram. Finally, it explains how the participant can use this information to decide which is better (see Figure 7). This procedure is roughly equivalent to graphically estimating a standardized score.

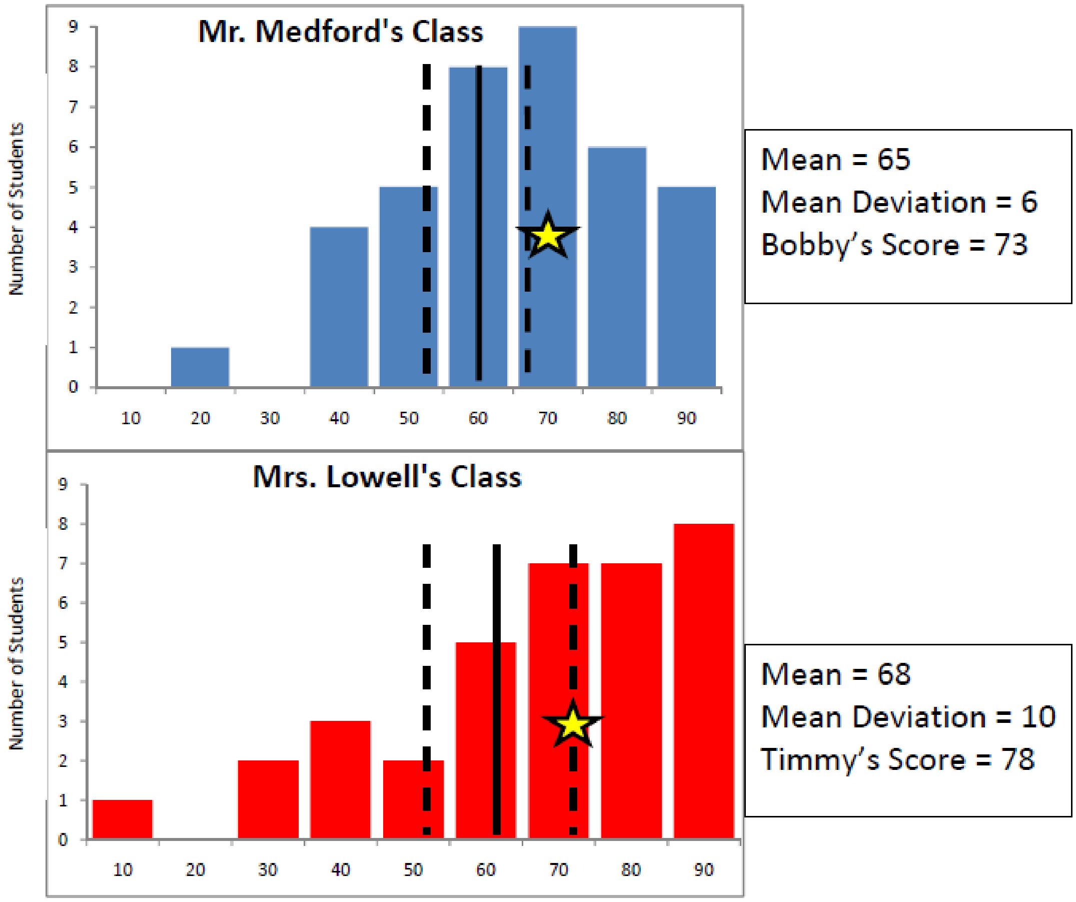

Figure 7. Worked example on graphically estimating a standardized score.

Immediately following this worked example was another learning activity problem ("Track Stars"). In this problem, students are asked to decide which of two world records, from two different track and field events, was "more shattered." Students are given a set of scores from two different events, and two exceptional values for each (one for Bill, one for Joe), and told to use math to help them decide whether Bill or Joe had a more impressive performance, given the rest of the competitors (see Figure 8). Both the "Pitching Machines" and "Track Stars" problems were coded following a zero, one, two scheme as well, with zero representing no work or a totally incorrect answer, a one representing an answer that used a procedure that reflected some aspect of variability, and a two representing a correct answer (arrived at either mathematically or graphically).

Track Stars

Bill and Joe are both on the U.S. Track Team. They also both broke world records last year. Bill broke the world record for the high jump with a jump of 8 ft (96 inches). Joe broke the world record for the long jump with a jump of 26 ft (312 inches). Now Bill and Joe are having an argument. Each of them thinks that his record is the best one. You need to help them decide. Based on the data below, decide if 96 inches shattered the high jump record more than 312 in. shattered the long jump record.

Use this page and the next to make any calculations and draw any graphs you need. Please write your answer at the bottom of the next page.

Top High Jumps in 2000		Top Long Jumps in 2000	
Height (inches)	**Number of Jumps**	**Height (inches)**	**Number of Jumps**
78	1	258	1
80	2	264	2
82	3	270	2
84	5	276	9
86	6	282	9
88	7	286	4
90	4	292	1
92	1	296	1
96 (Bill)	1	312 (Joe)	1

Figure 8. The "Track Stars" problem, which asked participants to decide which of two records was more exceptional.

2.1.3.3 Activity Questionnaires

Following the learning activities, participants completed two pages of questionnaires. The first page was the manipulation check, and the second measured their task-based goal adoption during the learning activity. The manipulation check asked students "At the beginning of the learning phase, you were asked to focus on just one goal for this study. What was it?" and told them to check only one of the possible responses. These responses corresponded to performance-avoidance, performance-approach, and mastery-approach, respectively; specifically, the items read "My goal was to not perform poorly compared to other participants," "My goal was to perform well compared to other participants," and "My goal in this study was to develop my understanding of these materials." This was modeled on the manipulation checks used in prior research (Elliot et al., 2005).

The second page was a measure of their achievement goal adoption during the just-completed learning phase. Specifically, this questionnaire had seven items (see Table 3) which were designed to assess certain features of mastery-approach goals (skill development and improved understanding), performance-approach goals (doing better than others), and performance-avoidance goals (not doing worse than others). One item asked about a non-specific approach goal ("While completing these activities, I focused on answering everything correctly"), and is not included in further analyses. Additionally, there was a forced-choice question, which asked, "While completing these activities, which were you most trying to achieve?" and required the participants to choose one of two options, "Trying to understand" or "Trying to answer correctly." This was designed to measure the degree to which a mastery or performance goal had been adopted in the various conditions.

Table 3

Activity Questionnaire Items

Construct	Questions	
Mastery-Approach	*While completing these activities, I tried to improve my skill in solving these kinds of problems.*	*While completing these activities, I was focused on improving my understanding.*
Performance-Approach	*While completing these activities, I tried to perform better than other students.*	*While completing these activities, my main focus was to outdo other students.*
Performance-Avoidance	*While completing these activities, I was concerned that other students were performing better than me.*	*While completing these activities, I was concerned that I was making mistakes that others were not.*

2.1.3.4 Post-Test

The post-test included, in order, three procedural fluency items, a worked example on standardization, a mean deviation word problem, an open-ended explanation problem, and a transfer problem. As in the pre-test, the procedural fluency questions provided data sets for participants to work with and asked them to calculate the mean and mean deviation for these data sets. Different difficulty levels were created, with data sets consisting of small single- and double- digit numbers, another with larger double-digit numbers, and another with three-digit numbers.

The worked example was the same as used in prior research (Belenky & Nokes-Malach, 2012; Schwartz & Martin, 2004). It first describes a scenario (with data) in which a standardized score was needed to help determine which of two performances in different events was better, demonstrates the formula for a calculating a standardized score (see Figure 9), and instantiates that calculation for the situation described. This is followed by another very simple problem and a prompt to use the formula to solve it, which all participants solved correctly.

The mean deviation word problem provided an opportunity to see if students understood the procedures at a deep or surface level. The problem described groups of scores that students in a class had received in the following way:

> *Twenty students took a midterm in their science class, and they had an average score of 75. Five of them scored 70, five students scored 65, five students scored 80, and five students scored 85. What is the mean deviation?*

Students could answer this problem by doing all of the math out (that is, counting five scores at each of the four values), or, if they recognized the fact that there were five of each score meant that each of those values is weighted identically, they could simply find the mean deviation of

Standardization:

A standardized score helps us compare different things. For example, in a swim meet, Cheryl's best high dive score was an 8.3 and her best low dive was a 6.4. She wants to know if she did better at the high dive or the low dive. To find this out, we can look at the scores of the other divers and calculate a standardized score (see Table C1).

To calculate a standardized score, we find the average and the mean deviation of the scores. The average tells us what the typical score is, and the mean deviation tells us how much the scores varied across the divers. Table C2 presents the average and mean deviation values.

The formula for finding Cheryl's standardized score is her score minus the average, divided by the mean deviation. We can write:

$$\frac{\text{Cheryl's Score} - \text{average score}}{\text{Mean deviation}} \quad \text{or} \quad \frac{X - M \text{ of } x}{M \text{ deviation of } x}.$$

Table C1

Diver	High Dive	Low Dive
Cheryl	8.3	6.4
Julie	6.3	7.9
Celina	5.8	8.8
Rose	9	5.1
Sarah	7.2	4.3
Jessica	2.5	2.2
Eva	9.6	9.6
Lisa	8	6.1
Teniqua	7.1	5.3
Aisha	3.2	3.4

Figure 9. The worked example on calculating standardized scores, which was embedded in the post-test.

the four values. Performance on these problems were scored trichotomously, either incorrect (zero), correct but inefficient (one), or correct and efficient (two), based on the calculations made to solve the problem.

The open-ended explanation asked the participant to explain the mean deviation procedure to a naïve friend who does not know anything about it. Specifically, it asks:

> *Your friend who has never calculated mean deviation before asks you to explain the procedure to him. What would you tell your friend so that he could understand how to calculate mean deviation?*

This problem was intended to be one that indicated the depth of conceptual understanding that participants had attained. Responses were coded for particular features of the procedure, which were used to score the response as a zero (incorrect or not attempted), one (an answer which is on the right track but would not produce the mean deviation formula) or two (a description which could be followed to compute the mean deviation correctly). The difference between a one or two frequently was decided by whether the response explicitly mentioned taking the absolute value of a data point's deviation; without this adjustment, the formula does not produce the correct answer.

The final problem participants answered was the post-test transfer item. This word problem was isomorphic to the pre-test measure. The cover story for this problem had to do with teenagers arguing about their scores on a driver's license test, which they took with different teachers. As this problem was isomorphic to the pre-test measure, students might create a number of incorrect responses based on particular misconceptions (i.e., higher number is best, range from "mean plus deviation" value). The same zero (incorrect), one (conceptually correct), two (computationally and conceptually correct) coding scheme was used for this problem as had been used for the pre-test transfer problem.

2.1.3.5 End-of-Experiment Surveys

The questionnaires administered at the end of the experiment assessed participants' task-based goal and strategy adoption during the experiment, their course-based achievement goals for mathematics, and demographic information. The first questionnaire asked students to write a freeform response to the prompt, "What was your goal in this experiment? That is, please explain what you focused on achieving during the experiment." The next questionnaire was a modified version of the manipulation check used earlier. Specifically, this questionnaire provided three goal statements, and asked them to respond to those statements according to a seven-point Likert scales, ranging from "Strongly Disagree" (1) to "Strongly Agree" (7). Formatting the questionnaire in this way allows for the measurement of multiple goal adoption. The specific wording of the performance-avoidance, performance-approach, and mastery-approach items, respectively, were “My goal was to perform better than other participants in this study,” “My goal was to avoid performing worse than other participants in this study,” and “My goal in this study was to develop my understanding and skill in solving these types of math problems.”

The following page assessed students’ self-reported strategy usage, in two forms. The top half of the page asked students to “Think back to how you went through both the learning and test sections of the experiment. Did you do any of the following things? (Check all that apply).” After that was a list of six choices that students could respond to, which were: 1) I checked my responses to see if they made sense; 2) I checked my responses to see if I made any errors; 3) I explained what was presented in the learning materials to myself; 4) I made a plan for how to solve a problem before beginning to work on it; 5) I thought of things from my own life that the problems were similar to; 6) I tried to notice similarities between the problems that were presented. The bottom half of the page had space for students to write a freeform response to the prompt “Did you use any strategies during this experiment? Think back to both the learning and test sections, and write down anything you consciously did or noticed yourself doing.” A rubric was developed to code this freeform response, based on an iterative process of generating a set of strategies we expected to see, analyzing what sorts of responses were common in a subset of participants’ data, and then combining these into a final rubric. This rubric contained a list of reading strategies (i.e., reading carefully, underlining, taking notes, skimming for important information), problem-solving strategies (i.e., follow instructions, follow examples, outline problem-solving steps, find analogies for test problems from learning materials, math-specific strategies, write down their work), and learning strategies (i.e., memorizing, focusing on understanding concepts), as well as a few other miscellaneous comments which occurred somewhat frequently (i.e., relying on prior experience, answering as many as possible, concentrating on talking aloud, or none). As there were a number of reading and problem-solving strategies, these were summed to form a composite score for use in the analyses.

Following this page was the AGQ-R (Elliot & Murayama, 2008). This instrument presents twelve prompts, with three prompts corresponding to each of the four achievement goals (mastery-approach, performance-approach, mastery-avoidance, performance-avoidance). The prompts were preceded by instructions, which asked participants to focus on their current math class (or, if they were not currently taking a math class, to think of their most recent one) when responding to the prompts. These responses were assessed using a 7-point Likert scale, ranging from “Strongly Disagree” (1) to “Strongly Agree,” (7) with higher scores indicating greater endorsement of the particular achievement goal, following established scoring conventions (Elliot & Murayama, 2008). This is a revised version of one of the most widely used measures for assessing course-level achievement goals (the Achievement Goal Questionnaire, or AGQ; Elliot & McGregor, 2001). While prior research using similar materials (Belenky & Nokes-Malach, 2012; Belenky & Nokes-Malach, in press) had administered the AGQ at both the beginning and the end of the experiment, the current study only administered the AGQ-R at the end. This decision was made for a number of reasons, based on prior research, theory, and due to pragmatic concerns (i.e., eliminating potential confounds in the experimental design and keeping the experiment within a certain short timeframe). The same prior research gives no reason to believe that responses to the AGQ-R, which assess course-level achievement goals, are influenced by momentary changes in task-based achievement goals. Other research from social and cognitive psychology on priming indicates that students can be influenced by goal-related stimuli (e.g., Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001). In particular, these studies have shown that presentation of goal-relevant words can trigger unconscious goal activation, which can in turn influence goal-driven behaviors. Asking participants to fill out the AGQ-R before the experiment may have had similar effects.

The theory of achievement goal motivation would also predict that being reminded of one's course-level goals would make one more likely to be influenced by those goals in a task immediately following the reminding (Elliot, 1999). Perhaps more critically, because the experiment called for a manipulation of achievement goals, avoiding reminding participants of their existing goals helps maintain internal validity. Any differences observed in goal adoption, learning behavior, and transfer outcomes should be due exclusively to the manipulations, and not due to reminders of existing goals, or interactions between that reminding and goal manipulations. Finally, pragmatically, the experiment was designed to be completed in one hour or less, and minimizing the amount of time answering questionnaires helped to achieve that goal.

2.1.4 Procedure

The study consisted of a series of discrete phases, most of which corresponded to clearly demarcated sections of the binder that participants went through within given time limits. Specifically, subsequent to the informed consent process, participants completed, in order, the following phases: Pre-Test (5 minutes), Talk-Aloud Training (~3 minutes), Goal Manipulation (~3 minutes), Learning Activities (20 minutes), Activity Questionnaires and Post-Test (20 minutes), End-of-Experiment Surveys (~ 5 minutes), Debriefing (~3 minutes). Figure 9 presents a graphical representation of this procedure.

Most elements of the study were instantiated as physical materials through the binder, as was discussed in section 2.1.3. The aspects that were managed directly by the experimenter (the talk-aloud training, the goal manipulation, and the debriefing) will be discussed next.

2.1.4.1 Talk-aloud Training

As part of the informed consent process, participants were made aware that they would be video- and audio-taped during the experiment. After completing the pre-test, the experimenter delivered the following text from a script:

> *"So, before we move onto the second section, let me give you a few more instructions for what I'd like you to do in this experiment. As you go through the rest of the experiment, there are going to be parts where I ask you to talk aloud, verbalizing whatever you are thinking. That is, I just want you to say out loud whatever is in your mind at that moment. It is not doing any extra thinking, or different thinking, but just saying whatever it is you are thinking. So, to practice that, I will give you some simple arithmetic problems; try solving them out loud to practice."*

Participants were then handed a sheet of paper that had four simple arithmetic problems on them. If participants did not talk aloud as solving them, or struggled to talk aloud in the correct way (that is, without reflection, but simply saying what was in their working memory at the time; Ericsson & Simon, 1993), the experimenter modeled one way to solve the first problem while talking aloud. When participants completed all four of the problems and the experimenter was satisfied that they understood the idea of how to talk-aloud, the experimenter took the sheet away from the participant and moved onto the goal manipulation.

2.1.4.2 Goal Manipulations

Goal manipulations were constructed based upon a close reading of prior research on experimental manipulations of achievement goals (e.g., Elliot & Harackiewicz, 1996, Senko &

Harackiewicz, 2005). A set of manipulations were created that framed the purpose of the materials, and explicitly provided a goal that the participant should focus on. Piloting of this material was conducted through iterations of a “rapid prototyping” cycle, where a group of participants would use the current versions of the materials, manipulation check results would be evaluated, and revisions would be made and then tested with a new group of participants, and so on. The changes that were made included additional information about the framing, as well as attempts at illustrating for participants exactly what sort of performance they should aim for in the performance-approach and performance-avoidance condition. In particular, some versions included a histogram of (fabricated) performance data, and told participants where on the histogram they should aim to score better than (or no worse than, in the performance-avoidance condition). For most of the pilot studies, these manipulations were delivered as text to the participants, on the front page of their packet, but, ultimately, this appeared to be creating weaker manipulations than when the framing instructions were delivered by the experimenter. In total, 281 participants went through one of three versions of the mastery-approach manipulations, or one of the nine versions of the performance-approach and performance-avoidance materials, before the final materials were created.

Goal manipulations in the present study reported here were delivered verbally by the experimenter, ostensibly as additional information about the purposes of the experiment. Specifically, immediately following the talk-aloud practice, the experimenter put down the script and struck a more conversational tone with the participant, saying “So, before we go on, I wanted to talk you a bit more about the experiment. Could you tell me, is this your first one of these psychology experiments this semester?” Depending on the response, the experimenter then added some variation of the statement “Well, sometimes when you do these psych experiments, you can kind of feel like ‘What is going on here? Why are they having me do this?’ I find that it works better when participants have a sense of what the study is all about, so I’m going to tell you a bit more about the thinking behind the study.”

This was followed, in each condition, with a common introductory section. This described, in general terms, different goals that students may have, before moving onto the specific goal that the participant was to focus on for the rest of the experiment. Specifically, the text of this general introduction stated:

> *So, you know how people can have different motives or goals, depending on the situation? Like, for example, sometimes people focus on demonstrating how much they can do, and how good they are at it. Or, other times, they just want to not do poorly compared to everyone else. Or, other times, they are trying to improve their own abilities or understanding. These are all perfectly fine goals that may match the situation, but in this study, I really want you to focus in on just one goal...*

This was immediately followed by a statement of the goal they should adopt, depending on their condition. The mastery-approach condition was told that their goal was “to develop your understanding of these materials and your skill in solving these types of problems.” The performance-approach condition was told that their goal was “to perform well compared to other participants,” while the performance-avoidance condition was told that it was “to not perform poorly compared to other participants.”

After these instructions, participants received more specific instructions, depending on their condition. These instructions described more fully the purpose of the study, how participants should consider evaluating their competence as they go through the study, a reason for videotaping and having them talk-aloud, and, finally, information about the type of feedback information they will receive at the end of the study.

Participants in the mastery-approach condition were told:

"Specifically, we are interested in developing a set of materials that help students learn this material well. The intention is to make learning materials that aid Pitt students in improving their understanding of basic mathematical concepts. So, I really want you to focus on doing just that; try to develop your understanding of these materials, ok? As I said, I'm going to record you; this is so that we can look at how you improve your learning as you go through these materials. At the end of the study, I will take your packet and be able to give you feedback on how you much you improved from the beginning of the study to the end."

Participants in the performance-approach condition were told:

"Specifically, we are videotaping a lot of students going through these materials, and are trying to find those that produce better performance than most of the other participants. We know this material results in a wide range of performance, which means that sometimes people perform at an average level while others perform exceptionally well. We want to identify those people, and specifically, look for instances when people do particularly well on any given part of this experiment when compared to other students in this experiment. Then, we plan on showing the videos of really excellent performances to other students, so that they can see what an excellent performance looks like. So, as much as you can, I want you to focus on trying to perform better than the majority of other participants throughout the study, ok; really stand out, you know? At the end of the study, I will take your packet and be able to give you feedback on how you performed relative to other participants, and whether or not your video may be one of the ones that are used to showcase high level performance on these materials."

Participants in the performance-avoidance condition were told:

"Specifically, we are videotaping a lot of students going through these materials to examine when people perform well or poorly. We know this material results in a wide range of performance, which means that sometimes people perform at average level while others perform exceptionally poorly. We will examine each person's performance and compare it to other students to find instances when people do particularly poorly on any given part of this experiment. We will use these videos of poor performance to show other students examples of what not to do when solving these problems. So, your goal is to try and ***not*** *perform any worse than the majority of other participants throughout the study; don't stand out for doing too poorly, you know? At the end of the study, I will take your packet and be able to give you feedback on how you performed relative to other participants, and whether or not your video may be one of the ones that are used to show exceptionally poor performance."*

In each condition, these instructions were followed by the experimenter asking the participant to repeat back what their goal for the study should be. All participants in the mastery-approach condition were able to do so. Participants in the performance-approach and performance-avoidance condition, however, sometimes responded in ways that did not match the goals they were expected to say. Specifically, of the 35 participants in the performance-approach condition, three gave first responses that were not clearly performance-approach goals, usually responding with a somewhat general, mastery-type goal (e.g., “I will try to do my best and see how well I can do”). Of the 35 performance-avoidance participants, ten gave first responses that were better described as performance-approach goals. In all of these cases, the experimenter made clear that they should really focus on the particular goal, based on their condition (“To perform better than the majority of other participants” or “To not perform any worse than the majority of other participants”).

2.1.4.3 Debriefing

Participants were debriefed about the purposes of the study upon completing the end-of-experiment surveys. This debriefing included a discussion of the different types of achievement goals that researchers have studied, and what effects they may have on learning behaviors, affect, and achievement outcomes. Participants were also given a debriefing sheet that described the hypotheses of the current study, provided references to published research on achievement goals, and included contact information that the participant could use to learn more information or ask further questions they might have. In addition, they were given a blank copy of the consent form for their records, which also contained contact information for the experimenter and the University of Pittsburgh Human Subjects Protection Advocate. The experimenter provided an opportunity for the participant to ask any questions they may have still had.

2.1.5 Talk Aloud Protocols

Talk-aloud protocols were recorded using the built-in webcam on an iMac and EvoCam [computer software]. Audio was recorded through an externally connected Blue Snowflake™ microphone. These videos, which continuously recorded both the learning and test phase, were transcribed by the author and several research assistants. The protocols were broken down into utterances, defined as one classifiable thought, usually at the level of a sentence. An utterance lasted anywhere between one and ten seconds when participants were working on problems. When participants were reading text or making long calculations, longer streams of text were coded into one utterance. Periods of silence longer than a few seconds typically indicated the beginning of a new utterance. The length of these utterances was recorded, as well as whether or not the student was referring to a prior example or the current problem being worked on.

To construct the rubric for coding utterances, a specific set of cognitive strategies and processes were identified, based on analysis of materials used in the study and past research that coded talk-aloud protocols of learning (Chi et al., 1989; Gadgil, Nokes-Malach & Chi, 2012; Renkl, 1997). Self-explanation was included, defined as an inference beyond the provided text or problem to provide underlying justifications or reasons (see Table 4). Other beneficial cognitive processes that were expected to be used were analogical comparison between problems, qualitative reasoning about aspects of the problem and the procedures used to solve it, as well as error correction (both at a conceptual level, and that the level of correcting errors in calculations). Additionally, as prior research had observed a stark difference in the amount of comprehension monitoring between those who learned well and those who did not, these were coded as well. These coding categories formed the basis for the original rubric that was used.

A process of iterated revision was used to refine this rubric to more closely match the data. Specifically, the original rubric was used by the author to classify a subset of data. During that coding, a number of other types of elaborative statements were noted, and the rubric was subsequently revised to more accurately reflect all of the types of utterances made by participants. In particular, additional categories were added (e.g., within-problem comparisons, descriptions of problem-solving actions, etc.) to more accurately capture the range of constructive statements participants made. This revised rubric was used by the author and a research assistant on a subset of the collected protocols. Disagreements between the two coders were discussed. Many instances could be resolved, with one coder agreeing that a different category was the more appropriate one. In cases where the discussion revealed that there were differences in how the coding categories were interpreted, or where they were unclear, or where they overlapped too closely with another category, the definition was revised to address these issues. The final coding rubric (see Table 4) was then used by the author to code each of the utterances recorded in the protocols.

Table 4

Rubric Used to Code Elaborative Statements

Code Type	Definition	Example statement
Explanation	Inferring beyond the provided text or problem to provide underlying justifications or reasons.	*"The sample size is the same so you don't have to account for that"*
Between Problem Comparison	Drawing a comparison to another problem	*"So what we can do was, uhm, the same thing for the previous pages."*
Comparing Test Material to Learning	Drawing a comparison on the transfer problem to the worked example, or earlier problems	*"So we'll just um, do that procedure we were doing before. Like the standardized number."*
Qualitative Reasoning	An estimation (about the outcome of a problem, or a particular step) made without doing the work out.	*"The fireball… looks the most, looks the most accurate."*
Evaluation	A question or some statement that indicates the intent to evaluate whether one is on the right track, or to check their work.	*"Let's just go back and check."*
Conceptual Error Correction	A statement that reflects an understanding that something is incorrect with their strategy or reflect noticing a misconception about the problem.	*"Oh yeah then divide by 5 because there's 5 values."*
Calculation Error Correction	Noticing of a small error that is not explicitly conceptual. Small calculator errors would fall into this category	*".364, 1.346, no. 1.364, minus 2…"*
Question	An open-ended question to oneself or a statement that some particular problem step or feature is unclear.	*"But how do they get the deviation?*
Within-Problem Comparison	Contrasting and/or comparing within a problem.	*"Yeah, Susan got a higher test score with the test with a lower average."*
Describing a Feature	Describing some feature of the problem, that has been noticed or observed.	*"So Susan scored above the dev… one deviation above the mean."*
Describing Action	The description of a problem-solving step being taken or about to be undertaken.	*"Starting at 258, 296 is about 40, we'll go in increments of 10."*
Negative Monitoring	A statement which reflects a lack of understanding, in general.	*"I'm not sure some of my thinking was that I don't know how well I did on this."*
Positive Monitoring	A statement which reflects acknowledgement of understanding, in general.	*"Well, now I think I'm fine."*
Negative Affect	A statement which reflects a negative emotional or affective state.	*"I hate these."*
Positive Affect	A statement which reflects a positive emotional or affective state.	*"Interesting, uhm"*

3.0 RESULTS

The hypotheses being tested in this study were that 1) mastery-approach goals would lead to higher rates of transfer 2) mastery-approach goals would lead to increased self-explanation, and 3) higher levels of self-explanation would predict greater transfer. Before testing these hypotheses, it is important to first establish that the manipulations themselves were successful, and produced a pattern of goal adoption that indicates that students were differentiating between these goals.

3.1 GOAL ADOPTION

In total, there were four measures of goal adoption in the task; a forced-choice manipulation check after the learning phase, an activity questionnaire immediately after that, an open-ended prompt administered after the test phase about what goals a student adopted in the experiment, and a three-item Likert-scale questionnaire that followed it. These will be addressed in that order.

The forced-choice manipulation check administered after the learning phase clearly showed that, overall, participants recalled which goal they were asked to focus on, χ^2 (4, N = 105) = 179.35, p <.001, with 94% of the participants correctly choosing their condition.

This was followed by the activity questionnaire, which asked students to respond via a Likert-scale about their goals during the recently completed learning phase. Two items for each goal construct were administered, and these were adequately reliable (Cronbach's α = .61 for MAP, .84 for PAP, .88 for PAV). As such, construct scores were created by averaging responses across the two values, creating an activity questionnaire score for each construct. A series of planned comparisons were made between each condition on each activity questionnaire score.

Examining each construct score separately, activity mastery-approach goals were highest for the mastery-approach condition, compared to either the performance-approach, t (68) = 2.59, p = .012, d = .31 or performance-avoidance conditions, t (68) = 4.13, p < .001, d = .50 (see Figure 10). Activity performance-approach goals were higher for the performance-approach condition, compared to the mastery-approach, t (68) = 9.35, p < .001, d = 1.13, and marginally higher than for the performance-avoidance conditions, t (68) = 1.95, p = .056, d = .23. Activity performance-avoidance goals were higher for the performance-avoidance condition compared to the mastery-approach condition, t (68) = 5.16, p < .001, d = .62. However, there was no difference between the performance-avoidance and performance-approach conditions, t (68) = .07, p = .943. Each experimental condition produced (or was tied for) the highest ratings for the goal that corresponded to the condition, providing evidence that the manipulation was successful in triggering the appropriate dominant goal in the participant. That said, the performance-approach condition led to as much performance-avoidance goal adoption as the performance-avoidance condition.

The activity questionnaire also included a forced choice item that asked students which of two options they were most trying to accomplish during the just-concluded learning phase; to understand, or to answer correctly. There were significant differences between the mastery-approach and the two performance goal conditions, χ^2 (2, N = 105) = 31.37, p <.001. The majority of the mastery-approach condition (66%) reported trying to understand as their stated goal, while almost all of the performance-approach (83%) and performance-avoidance (91%) condition reported trying to answer correctly as their stated goal. This provides additional evidence that participants focused on the goal that the manipulation was targeting.

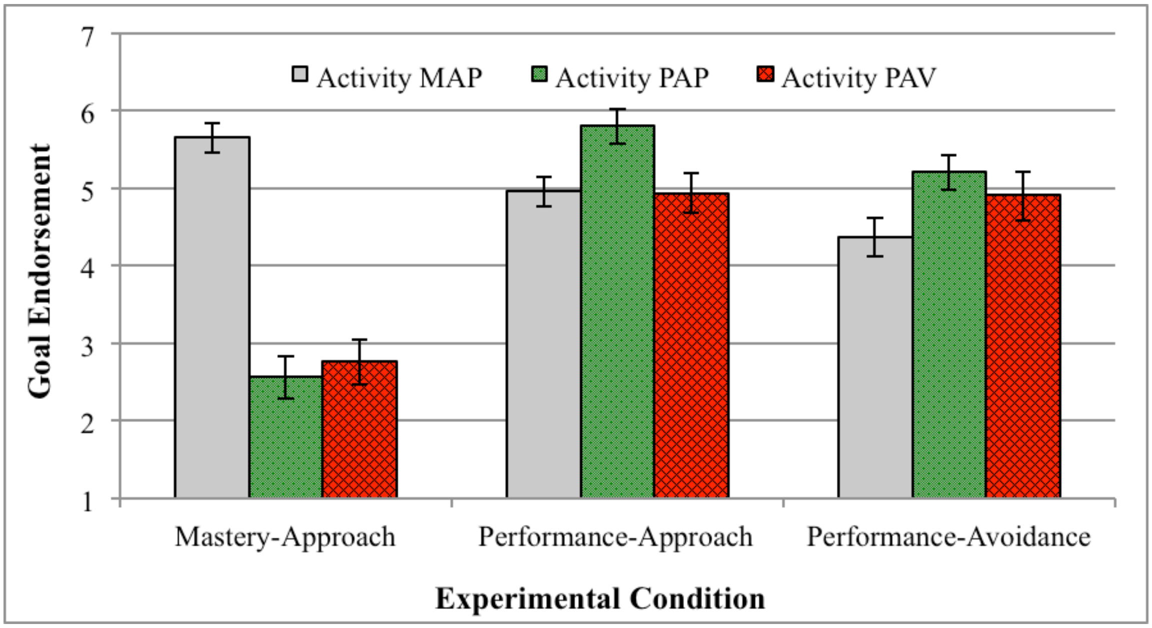

Figure 10. Goal endorsement on the activity questionnaire for each condition. Error bars represent standard errors.
Note. MAP = Mastery-Approach, PAP = Performance-Approach, PAV = Performance-Avoidance.

Additional measures of goal adoption were assessed during the questionnaire phase, at the end of the study. The first such measure in the questionnaire phase asked students to respond to the open-ended prompt, "What was your goal in this experiment? That is, please explain what you focused on achieving during the experiment?" Participants spontaneously generated a response which matched their experimental condition 81% of the time. For example, mastery-approach goal students were likely to respond with answers such as "To understand the material that was presented to me and try to get better at what I learned." Performance-approach goal students tended to produce statements such as "My goal was to do better than other students." The performance-avoidance goal condition, in contrast, produced more statements such as "I tried not to do worse than the person doing the worst in the experiment." There were no significant differences in the accuracy rates of each condition, χ^2 (2, N = 105) = 3.09, p = .214. Other types of responses were also coded. There were no significant differences between conditions in the frequency of responses reflecting "doing one's best," χ^2 (2, N = 105) = 2.92, p = .233, "solving problems correctly," χ^2 (2, N = 105) = .42, p = .811, or simply to "get credit," χ^2 (2, N = 105) = 2.02, p = .364. However, participants in the performance-approach (49%) and performance-avoidance (51%) conditions were more likely to include statements of trying to get the correct answer than those in the mastery-approach condition (14%), χ^2 (2, N = 105) = 12.68, p = .002. Additionally, statements which reflected a goal of trying to finish the experiment were generated more frequently by the performance-avoidance condition (in 29% of responses), compared to the mastery-approach (3%) or performance-approach (9%) conditions, χ^2 (2, N = 105) = 11.04, p = .004. Even on an open-ended measure, there is strong evidence that the experimental manipulations led to a different pattern of goal adoption across the three conditions that is consistent with the goal constructs described by achievement goal theory.

The final measure of goal adoption participants completed was a set of Likert-scale questionnaires that allowed them to rate the degree to which they did or did not adopt each of the three relevant achievement goals, using similar wording to that used on the manipulation check administered earlier. The mastery-approach condition endorsed mastery-approach goals very highly (M = 6.31, SD = .90), while the performance-approach condition (M = 4.03, SD = 1.81) and performance-avoidance condition (M = 3.14, SD = 1.48), did not. There was a significant difference, revealed in a one-way ANOVA, F (2, 102) = 44.92, p < .001, η^2 = .47. Planned comparisons revealed that each of the conditions was significantly different from one another in the degree to which they adopted mastery-approach goals, ts (68) > 2.25, ps < .029. For performance-approach goal adoption, the performance-approach condition reported the highest levels (M = 6.31, SD = 1.13), compared to the mastery-approach (M = 2.11, SD = 1.57) and performance-avoidance (M = 4.83, SD = 1.64) conditions. These differences were significant, F (2, 102) = 74.26, p < .001, η^2 = .59, and planned comparisons again revealed that all three conditions were significantly different from one another, ts (68) > 4.42, ps < .001. Finally, for the performance-avoidance goal adoption measure, there were again significant differences, as evidenced in a one-way ANOVA, F (2, 102) = 45.99, p < .001, η^2 = .47. However, the performance-avoidance (M = 5.91, SD = 1.76) and performance-approach (M = 5.40, SD = 1.90) conditions were not significantly different from another, t (68) = 1.18, p = .243. Both the performance-approach and performance-avoidance conditions were significantly higher on this measure than the mastery-approach condition (M = 2.23, SD = 1.56), ts (68) > 7.65, ps <.001. Taken together, it appears that the mastery-approach condition cleanly produced a different pattern of responses than either of the other two conditions. While the performance-approach condition was higher than the performance-avoidance condition in adopted performance-approach goals, both performance goal conditions produced similar levels of performance-avoidance goal adoption.

To summarize, a relatively consistent pattern emerged across the different types of measurements. Participants in the mastery-approach condition endorsed a mastery-approach goal across multiple measures, while tending to not endorse either of the performance goals. The performance-approach condition led to strong endorsement of performance-approach goals, and much lower levels of mastery-approach. Students in the performance-avoidance condition very clearly reported having performance-avoidance goals during the study on both the forced-choice measure, and on an open-ended free response. However, on other measures that assessed goal adoption quantitatively, participants in this condition did not report higher levels of performance-avoidance goals than the performance-approach condition. It appears that the performance-approach manipulation seems to produce a degree of performance-avoidance goal adoption as well.

3.2 GOALS AND PERFORMANCE

3.2.1 Pre-Test Performance

The pre-test contained two items of interest. One required students to calculate the mean and mean deviation of a small set of numbers, and the other was an isomorph of the transfer problem in the post-test. The conditions had equivalent levels of performance on calculating the mean, with each condition having 86% (30 out of 35) generate the correct solution. There were no significant differences between conditions in correctly generating the mean deviation on the pre-test, χ^2 (2, N = 105) = 1.81, p = .405. There were, however, marginal differences between conditions on the pre-test transfer problem, χ^2 (4, N = 105) = 9.00, p =.061. As mentioned earlier, this problem was coded trichotomously, with a score of 0 reflecting incorrect answers, a score of 1 reflecting conceptually correct (i.e., the logic of standardized scores) but computationally incorrect responses, and a score of 2 reflecting a computationally correct answer (i.e., using standardized scores to choose the correct response). The differences indicate that the mastery-approach condition had more scores of 0 than the other two conditions on pre-test, prior to any experimental manipulations. Specifically, 86% of participants in the mastery-approach condition received a score of 0, while the performance-approach and performance-avoidance conditions had 57% and 63% of participants receive a 0, respectively. In addition, mastery-approach only had five scores of 1 or 2, whereas the performance-approach had 15 and performance-avoidance had 13. Subsequent analyses, which will be discussed in the following sections, were conducted with all of the participants included, as well as taking out those participants who scored higher than 0 on the pre-test transfer item. The pattern of results was the same in both sets of analyses.

3.2.2 Goal Conditions and Learning Phase Performance

Performance on the two learning phase problems was coded according to a trichotomous scoring scheme, with a zero representing an incorrect (or no) solution attempt problem, a one representing a procedure that was incorrect but reflected some aspect of variability, and a two representing a conceptually correct approach to solving the problem. For the pitching machines problem, finding the averages of the data sets and reasoning about those would earn a score of one, but calculating mean deviation and reasoning about those values would earn a score of two. For the track stars problem, a score of one is assigned if they attempted to reflect variability in some manner but without calculating, or graphically estimating, a standardized score.

There was a marginal effect of condition on the track stars problem, χ^2 (4, N = 105) = 9.00, p = .061. This was due to the performance-approach condition receiving marginally fewer scores of zero (9%) than the mastery-approach (23%), χ^2 (2, N = 70) = 5.53, p = .063, and the performance-avoidance condition (31%), χ^2 (2, N = 70) = 5.85, p = .054, with no significant difference between the mastery-approach and performance-avoidance conditions, χ^2 (2, N = 70)

= 2.10, p = .350. There was no effect of condition on the pitching machine problem, χ^2 (4, N = 105) = 4.18, p = .383. Altogether, these results indicate that the performance-approach condition produced slightly more sophisticated responses on one of the two learning activity problems than the other conditions.

3.2.3 Goal Conditions and Post-Test Performance

3.2.3.1 Goal Conditions and Post-Test Problems

A total post-test score was calculated by summing the score each participant received on the various problems in the post-test. Specifically, performance on the following problems was summed, with each being scored as a 0, 1, or 2, according to their coding; calculating mean and mean deviation of an easy data set (single-digit and low two-digit numbers), a medium data set (all two-digit numbers) and a harder data set (three-digit numbers), a mean deviation problem which could be solved efficiently or not, and explaining how to calculate mean deviation to a friend (the reliability of these measures was high, Cronbach's α = .87). This summed score reflects both procedural fluency in calculating, and the conceptual knowledge to describe how to calculate mean deviation. The only material on the post-test not included in this measure is the worked example on standardization and the transfer problem. The worked example on standardization is not included because all of the participants were able to correctly solve the simple problem provided with the worked example. The transfer problem will be analyzed separately.

This post-test score was used as a dependent measure in a one-way ANOVA, comparing the three experimental conditions. There was no significant difference in performance between the three groups, F (1, 102) = 1.81, p = .170. While the performance-avoidance condition (M =

7.49, *SD* = 2.95) scored higher than the performance-approach (*M* = 6.71, *SD* = 2.95) and mastery-approach (*M* = 6.26, *SD* = 3.14) conditions, this was a not statistically significant difference.

Subsequent analyses focused on each of the individual problems. There were no significant differences on the mean deviation calculations, *F*s (2, 102) < 2.12, *p*s > .128, the class test scores problem (which could be solved efficiently or not), *F* (2, 102) = .36, *p* =.697, or the "explain to a friend" problem, *F* (2, 102) = 1.67, *p* = .193[1]. The experimental conditions were equivalent in terms of their post-test performance.

3.2.3.2 Goal Conditions and Transfer Performance

The first main hypothesis of this study was that mastery-approach goals would produce higher levels of transfer. Previous research using similar test materials has used logistic regression models to predict the likelihood of successfully transferring, based on a set of predictor variables. That is, in these analyses the likelihood of a student receiving a score of two on the post-test transfer problem is modeled, based on some predictor variables. A logistic regression model predicting the likelihood of transfer based on the experimental condition was significant, χ^2 (2, *N* = 105) = 9.53, *p* = .009[2]. In these types of analyses, significant effects can be interpreted in terms of odds ratios – reported here as *Exp (B)* – that represent the change in degree of likelihood of the target outcome (in this case, successful transfer) for every unit change in the predictor. An odds ratio above one indicates how much higher the likelihood of transfer

[1]The same pattern of results (no significant differences between conditions on overall post-test performance or on individual post-test problems) was observed when entering pre-test transfer performance as a covariate in ANCOVA analyses.

[2]A separate analysis was also conducted taking out participants who had solved the problem correctly (i.e., scored a two) on the pre-test, as well as taking out those scored either a one (conceptually correct) or a two (computationally correct) on the pre-test. In both cases, the results followed the same pattern.

would be for every unit increase in the predictor, while a negative odds ratio would indicate the degree to which the likelihood is lower. This analysis revealed that the mastery-approach condition (49% correct) was less likely to transfer than the performance-avoidance condition (83%), Wald's χ^2 (1, N = 105) = 8.44, p = .004, *Exp (B)* = .20, but that there were no differences between the performance-approach (69%) and performance-avoidance conditions in the likelihood of successful transfer, Wald's χ^2 (1, N = 105) = 1.90, p = .169, *Exp (B)* = .45. The difference between the mastery-approach and performance-approach condition was only marginally significant, Wald's χ^2 (1, N = 105) = 2.84, p = .092, *Exp (B)* = .43. Contrary to predictions, mastery-approach goals did not increase the likelihood of transfer, relative to the other two conditions, but, rather, was associated with a lowered likelihood of transfer.

3.2.3.3 Adopted Goals and Transfer Performance

While the mastery-approach condition did not benefit transfer, it is possible that the degree to which a student adopts mastery-approach goals would increase the likelihood of transfer. To analyze this prediction, measures of student goal adoption in the experiment were entered as predictors of transfer in a logistic regression model. Specifically, a logistic regression was conducted which evaluated the likelihood of successfully solving the transfer problem with construct scores for mastery-approach, performance-approach, and performance-avoidance (taken from the activity questionnaire, see 3.1) as predictors.

This model was significantly different from a constant-only model, χ^2 (3, N = 105) = 9.78, p = .021. Within this model, the only variable which is significantly different from zero is the performance-approach construct score, Wald's χ^2 (1, N = 105) = 9.78, p = .002, *Exp (B)* =

1.52. For every unit increase in adopted performance-approach goals, the likelihood of transfer increased by 52%[3] (see Figure 11).

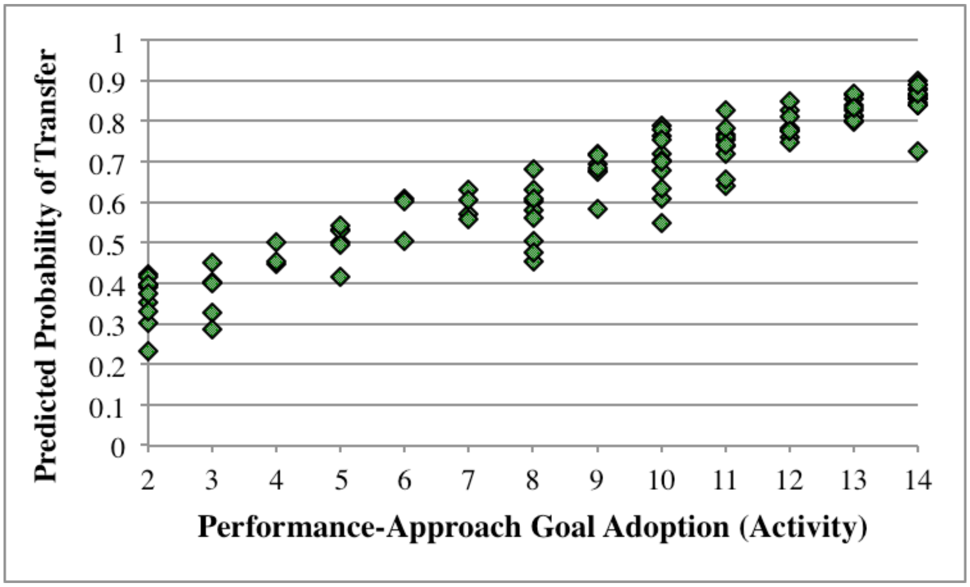

Figure 11. The relationship between performance-approach goal adoption, as measured by the activity questionnaire, and the predicted probability of transfer.
NOTE. The range for the x-axis is 2-14, as the construct score was calculated by summing the values from two 1-7 Likert-scale questions, with higher scores indicating greater goal adoption.

An additional measure of goal adoption came from the questionnaires at the end of the experiment. In one particular questionnaire, participants were asked to indicate the degree to which they had adopted each of the three achievement goals on a seven-point Likert-scale. These were entered in a separate model predicting the likelihood of transfer. This model was significantly different from a constant-only model χ^2 (3, N = 105) = 9.78, p = .021[4]. Within this model, the only variable which is significantly different from zero is the performance-approach item, Wald's χ^2 (1, N = 105) = 4.48, p = .034, *Exp (B)* = 1.30. For every unit increase in

[3] The same pattern of results was observed when limiting the sample to only those who scored a zero on pre-test, as well as when the sample was limited to those scoring either one or zero.
[4] Again, the same pattern was observed when limiting the sample in light of pre-test performance.

performance-approach goals reported on this questionnaire, the likelihood of transfer increased by 30% (see Figure 12).

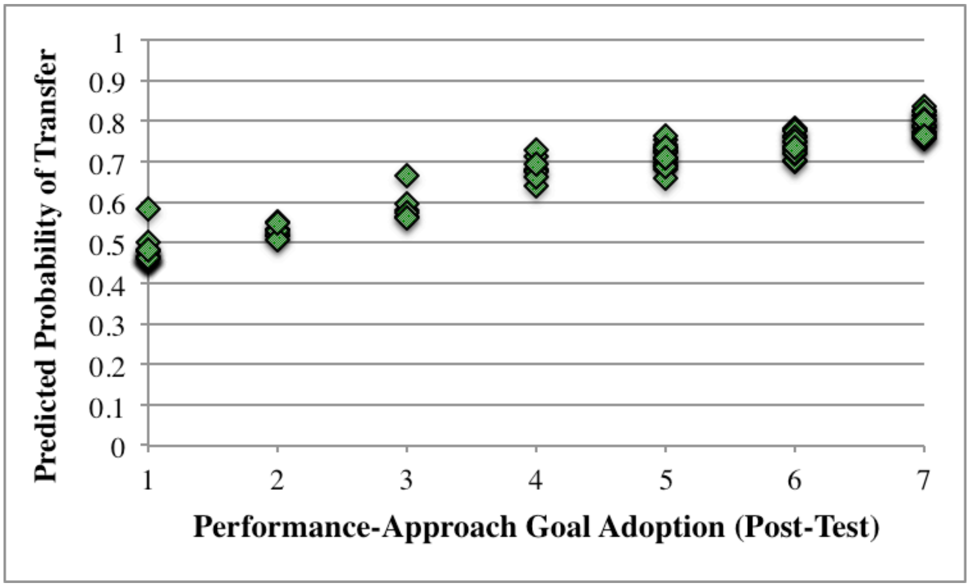

Figure 12. The relationship between performance-approach goal adoption, as measured by the post-test questionnaire, and the predicted probability of transfer.

In summary, the first hypothesis was that mastery-approach goals would lead to higher levels of transfer. This was not supported, either in terms of the experimental conditions, or in terms of self-reported goal adoption. Instead, the performance-avoidance condition produced the highest levels of transfer, compared to both the performance-approach and mastery-approach conditions. When analyzing the relationship between goals adopted during the experiment and transfer, adopted performance-approach goals (both measured after the learning phase and at the end of the study) were predictive of transfer, while the other adopted goals were not predictive.

3.3 GOALS AND PROTOCOLS

3.3.1 Goals and Self-Explanations

The second hypothesis of this study was that mastery-approach goals would lead to higher levels of self-explanations. Before testing this specific hypothesis, it is important to first examine the general patterns in the talk-aloud protocols across all of the transcribed components (i.e., learning phase, worked example, and transfer problem) of the experiment, by analyzing whether there were any between-group differences on the overall pattern of talk aloud data. There were no differences between the three conditions on the total number of statements made, the number of those statements coded as elaborative, nor on the amount of time spent on elaborative statements (see Table 5), all Fs (2, 101) < 1.27, ps > .287. Overall, participants made, on average 200.54 utterances (SD = 61.15) during the parts of the experiment which were coded (the learning phase, the standardization worked example, and the transfer problem), 33.49 of which were coded as elaboration statements (SD = 19.57). They spent an average of 139.01 seconds speaking these elaboration statements (SD = 83.85).

Table 5

Means (and Standard Deviations) of Talk-Aloud Behaviors, by Condition

	Mastery-Approach	Performance-Approach	Performance-Avoidance
Number of Statements	202.00 (70.20)	195.20 (49.57)	204.46 (63.57)
Number of Elaborations	34.38 (18.54)	32.17 (17.80)	33.94 (22.52)
Time on Elaborations (in seconds)	141.59 (74.65)	153.49 (103.52)	122.03 (68.22)
Number of Self-Explanations	3.62 (3.99)	3.66 (3.33)	3.57 (3.14)
Time on Self-Explanations (in seconds)	20.71 (22.91)	26.54 (23.79)	20.00 (19.20)

The most direct way to test the hypothesis that mastery-approach goals would lead to more self-explanation, is to examine the prevalence of these types of statements. Generally, it was quite low, (M = 3.62), although there was a substantial amount of variability (SD = 3.46). Specifically, 16% of participants generated no self-explanations, 61% of participants generated between one and five self-explanations, 18% generated between 6-10, and 4% generated 11 or more. Experimental condition was not predictive of the number of self-explanations produced across the whole experiment (self-explanations generated in each specific part of the experiment will be addressed in the subsequent section). Mastery-approach, performance-approach, and performance-avoidance produced similar levels of self-explanation, both in terms of the number of self-explanation statements, F (2, 101) = .01, p = .995, and time spent making self-explanation statements, F (2, 101) = .93, p = .40.

3.3.2 Goals and Other Types of Elaborative Statements

In addition to self-explanation, additional analyses investigated the other forms of elaborative statements that were coded for by the rubric (see section 2.1.5). Each type of elaborative statement was entered as a dependent variable in a one-way ANOVA, with the three experimental conditions as independent variables. Among the elaborations which focused on the content knowledge (as opposed to the student's own level of understanding or affective experience), only the number of times participants referenced past learning material (in particular, the worked example) during the transfer problem was significantly different between experimental conditions, F (2, 101) = 3.15, p = .047, η^2 = .059. A Bonferroni-corrected post-hoc analysis revealed that the performance-avoidance condition (M = .37, SD = .60) produced significantly more of these statements than the performance-approach condition (M = .06, SD = .24), but there were no differences between the mastery-approach condition (M = .26, SD = .67) and either of the other two. There was also a marginal effect of condition on the number of "describing features" statements participants made, F (2, 101) = 2.43, p = .093, η^2 = .046. In particular, the performance-approach condition (M = 3.54, SD = 2.74) was higher than the performance-avoidance condition (M = 2.26, SD = 2.44) in these types of statements; the mastery-approach condition (M = 2.88, SD = 2.09) did not differ from either of the other two conditions.

Some of the elaborations students generated reflected their own levels of understanding. Statements which reflected that a student believed that they understood or "got it" were labeled “positive monitoring” (as opposed to “negative monitoring” statements, which reflected a lack of comprehension). There were significant differences in the number of these positive monitoring statements made throughout the experiment, F (2, 101) = 4.04, p = .021, η^2 = .074. Bonferroni-corrected post-hoc analyses revealed that only the performance-approach (M = .14, SD = .43) and performance-avoidance (M = 1.03, SD = 2.12) conditions had significant differences, and that the performance-avoidance condition was higher on this measure. Additionally, there was a marginally significant difference on the number of positive affect statements, F (2, 101) = 2.42, p = .095, η^2 = .046. The mastery-approach condition (M = .32, SD = .95) was somewhat higher in these statements than either the performance-approach (M = .03, SD = .17) or performance-avoidance conditions (M = .09, SD = 37). All of the other types of elaborations were not significant (see Figures 13 and 14).

It is possible that that the goal adoption students reported was more predictive of elaboration statements than their experimental condition. As such, participants’ self-reported levels of goal adoption, as measured by the activity questionnaire construct scores (see Section 3.1), were entered as predictors of each type of elaboration statement in a series of linear regression models. The only one of these models which was significant predicted the number of negative monitoring statements, F (3, 100) = 3.01, p = .031. R^2 = .08. The only variable within this model which was significant was the performance-approach construct score, which was negatively related to the number of negative monitoring statements, B = -0.34, t(100) = 2.88, p = .005.

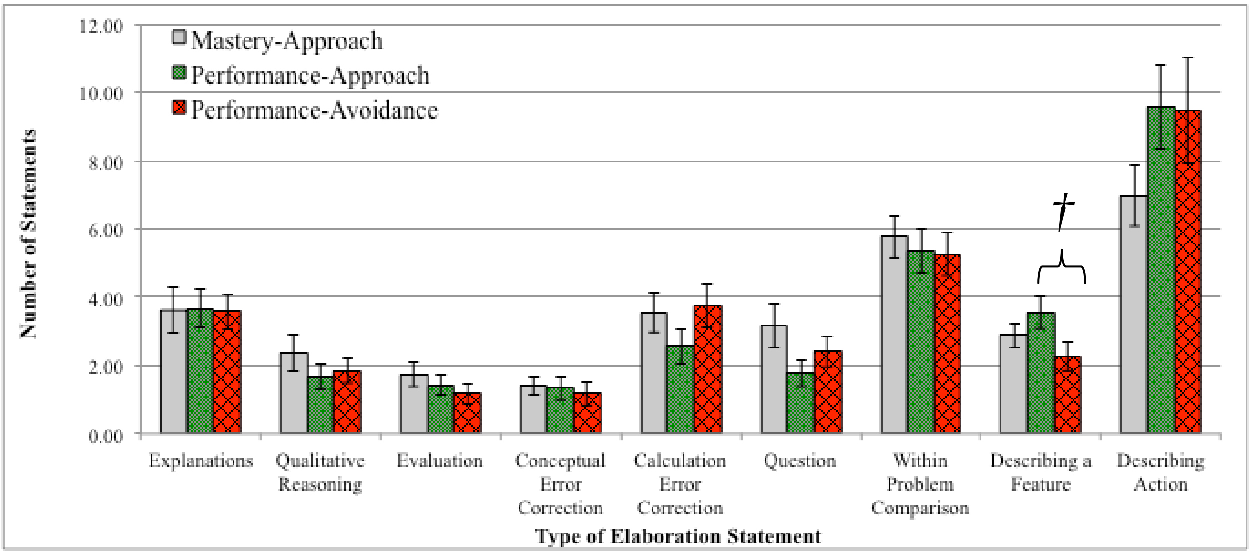

Figure 13. Mean number of total elaboration statements, across all coded utterances, for each condition, for statements with means above one. Error bars represent standard errors.
, p < .10.

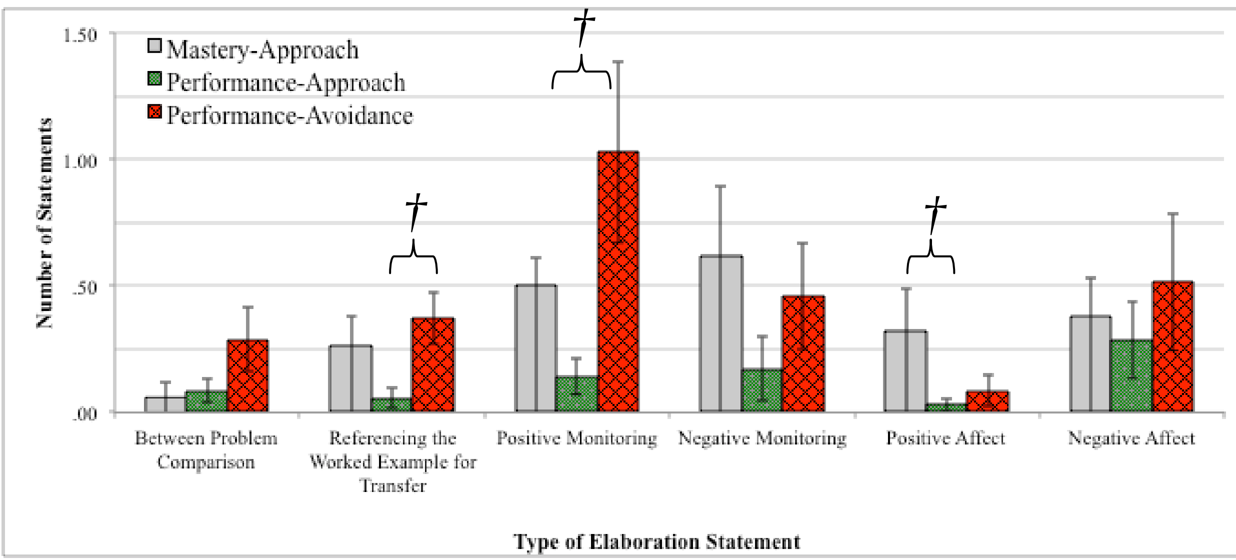

Figure 14. Mean number of total elaboration statements, across all coded utterances, for each condition, for statements with means above one. Error bars represent standard errors.
†, p < .10.

3.3.3 Talk-Aloud Differences for Each Section

The preceding analyses had looked at the differences in the amount of different types of elaborations throughout the parts of the experiment which were coded (the learning phase, the worked example on standardization, and the transfer problem). It is possible that differences in the amount of elaborations may have occurred on individual components of the experiment. To investigate this, the different types of elaborative statements made for each section of the experiment was analyzed separately, with the three experimental conditions entered as independent variables into a series of one-way ANOVAs.

3.3.3.1 Learning Phase Talk-Aloud Results

The four parts of the learning phase (Mean Deviation worked example, Pitching Machines problem, Graphical Standardized Score worked example, Track Stars Problem) were summed together. The number of positive monitoring statements was significantly different, F (2, 101) = 3.59, p = .031, with a Bonferroni-corrected post-hoc analysis showing that the performance-avoidance condition (M = .82, SD = 1.79) produced significantly more of these statements than the performance-approach condition (M = .11, SD = .40), but that there was no difference between the mastery-approach condition (M = .41, SD = .61) and either of the other two[5]. No other statistically significant differences were observed between the conditions in the

[5] Analyses were also conducted on each section of the learning phase separately. For number of statements, significant effects were only observed on the number of positive monitoring statements made during the track stars problem, with the same pattern as reported above; namely, that the performance-avoidance condition produced more of these statements than the performance-approach.

number of elaborative statements of each type (e.g., self-explanations, between-problem comparisons, etc.; see Table 5) during the learning phase, *F*s (2, 101) < 2.24, *p*s > .111.

For the standardization worked example, which was embedded in the post-test, the only significant difference between conditions was found for the number of self-explanation statements made, F (2, 101) = 4.00, p = .021. Post-hoc analysis revealed a higher degree of explanation statements for the performance-avoidance condition (M = .94, SD = 1.49) compared to the performance-approach condition (M = .37, SD = .54) and the mastery-approach condition (M = .38, SD = .49).

For the transfer problem, there were significant differences in the number of statements made overall, F (2, 101) = 3.17, p = .046, as well as the number of elaboration statements overall, F (2, 101) = 3.31, p = .041. For both of these, the mastery-approach condition (*M*s = 27.71, 7.50; *SDs* = 16.57, 6.42, respectively) was higher than the performance-approach condition (*M*s = 19.77, 4.71; *SDs* = 8.88, 3.19, respectively), but not statistically different from the performance-avoidance condition (*M*s = 22.06, 5.11; *SD*s = 13.84, 4.45, respectively).

In particular, there were significant differences on the transfer problem in the number of open-ended questions participants asked themselves, F (2, 101) = 3.36, p = .039, and the number of negative monitoring statements they made, F (2, 101) = 5.26, p = .007. For open-ended questions, post-hoc analyses revealed a difference between mastery-approach (M = .74, SD = 1.97) and performance-approach (M = .06, SD = .24), with a trend for mastery-approach to be higher than performance-avoidance (M = .14, SD = .55). For negative monitoring statements, mastery-approach (M = .18, SD = .38) was significantly higher than either performance-approach (M =.03, SD = .17), or performance-avoidance (which produced no negative monitoring statements). Students in the mastery-approach condition appeared to be more confused about how to proceed, and understood that they were struggling, more than in the other conditions.

Additionally, there were marginal effects for the number of comparisons between the transfer problem and the earlier worked example on standardization, F (2, 101) = 2.71, p = .071, as well as on the number of conceptual error correction statements made, F (2, 101) = 2.49, p = .088. Post-hoc analyses revealed that the performance-avoidance condition (M = .34, SD = .59) made more comparisons to the worked example than the performance-approach condition (M = .06, SD = .24), but that the mastery-approach condition (M = .26, SD = .67) did not differ from the other two. For conceptual error correction, post-hoc analyses revealed that the mastery-approach condition (M = .35, SD = .81) produced more of these statements than performance-approach condition (M = .09, SD = .28), but that the performance-avoidance condition (M = .11, SD = .40) did not differ from either of the two. Finally, there was a marginal effect of condition on the number of self-explanations made during the transfer problem, F (2, 101) = 2.39, p = .097. Bonferroni-corrected post-hoc analyses revealed that the mastery-approach condition (M = 1.24, SD = 1.44) made marginally more self-explanations during the transfer problem than the performance-avoidance condition (M = .66, SD = .77), but that there were no differences between the performance-approach condition (M = .89, SD = 1.02) and either of the other two.

3.3.4 Goals and Self-Reported Strategy Use

These analyses have focused on what was observed in the talk-aloud protocols. However, it may be instructive to consider whether the different experimental conditions self-reported using different learning strategies as they went through the materials. Participants had the opportunity to self-report the strategies they used during the experiment on a questionnaire administered at the end of the study. They were first asked to reply to a free-response prompt (see Figure 15), and then to check boxes from among a set of strategies (see Figure 16). The free response was coded for a set of features that had to do with reading strategies, problem-solving strategies, and

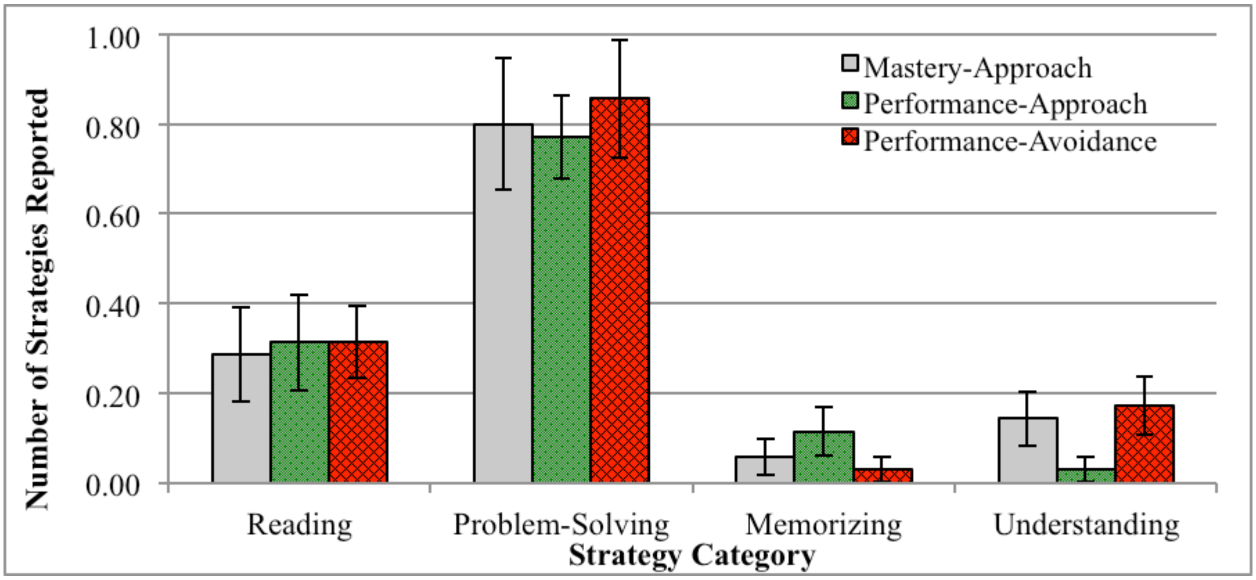

Figure 15. Mean number of self-reported strategies. Error bars represent standard errors.

learning strategies (specifically, focus on memorizing and a focus on understanding concepts). There were no significant difference between the conditions on the summed reading strategies value, F (2, 102) = .03, p = .972, nor on the summed solving strategies value, F (2,102) = .12, p = .886. There were also no significant differences between conditions on either of the two learning strategies; that is, there was no difference in the use of a memorizing strategy, χ^2 (2, N = 105) = 2.14, p = .343, nor the use of a focus on understanding strategy, χ^2 (2, N = 105) = 3.95, p = .139. Participants across the experimental conditions did not differ in how they responded to each of the six possible strategies that they could check, all χ^2s (2, N = 105) < 4.65, ps > .097.

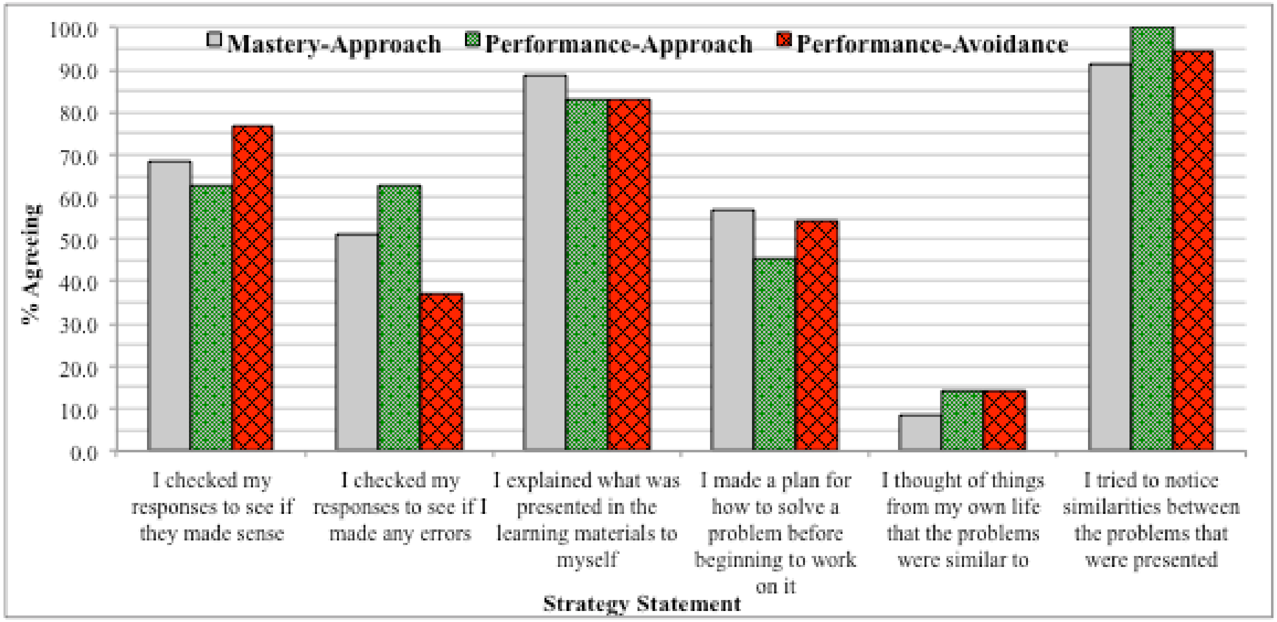

Figure 16. Percent of participants who endorsed using each strategy.

3.3.5 Summary

It had been expected that mastery-approach goals would lead to more self-explanation. However, this prediction was not supported. There were no differences on the number of self-explanation statements generated by the different experimental conditions, overall. However, the performance-avoidance condition generated more self-explanation statements during the worked example on standardization than either of the other two groups, contrary to expectations. The conditions differed on a small set of types of elaboration statements. In particular, the performance-avoidance and mastery-approach conditions produced higher levels of comparisons back to the worked example during the transfer problem. In addition, the performance-avoidance group reported more statements that reflected comprehension (i.e., "positive monitoring" statements) than the performance-approach condition. The mastery-approach condition did produce more elaborative statements on the transfer problem than the performance-approach condition; specifically, they asked themselves more open-ended questions and made more statements indicating a lack of comprehension. They also produced marginally more conceptual error corrections than the performance-approach condition, and marginally more self-explanations than the performance-avoidance condition. However, these differences in elaborations were not mirrored in students' own self-reports of strategy usage; on both an open-ended measure and a checklist measure, participants across all three conditions reported similar patterns of strategy usage.

3.4 ELABORATIONS AND TRANSFER

3.4.1 Explanations and Transfer

The third hypothesis being investigated was that self-explanations would produce higher levels of transfer, consistent with a large amount of prior literature. The number of explanations students generated overall were entered as predictors in a logistic regression model predicting the likelihood of transfer. This model was not significantly better than a constant-only model, χ^2 (1, N = 105) = .16, p =.692.

Additionally, explanations on each of the parts of the learning phase (i.e., mean deviation worked example, pitching machines problem, the graphical representation worked example, and the track stars problem) were entered as predictors into separate logistic regression models predicting transfer. Only the model which included explanations for the mean deviation worked example was a significantly better predictor of transfer than a constant-only model, χ^2 (1, N = 104) = 4.23, p = .040. Each of the five participants who produced an explanation statement on this worked example ultimately transferred successfully, making a coefficient representing the logarithmically-transformed odds for this variable unintelligible. The other models, which predicted transfer based on the number of explanations during the pitching machines problem, the worked example on graphically estimating a standardized score, and the track stars problem were not significant, χ^2s (1, N = 104) < .28, ps > .599. Models which predicted transfer based on the number of explanations generated on the worked example on standardization, embedded in the post-test, and during the transfer problem itself, were also not significant, χ^2s (1, N = 104) < 2.59, ps > .107.

The data was then restricted to each experimental condition to see if the relationship between explanations and transfer was different across the conditions. Within all three conditions, models predicting the likelihood of transfer based on the number of explanations were not significant, χ^2s (1, N = 34 for mastery-approach, 35 for performance-approach and performance-avoidance) < .69, ps > .406. When examining models predicting transfer based on explanations during learning restricting the data for each condition, no models were significantly better than a constant-only model, χ^2s (1, N = 34 for mastery-approach, 35 for performance-approach and performance-avoidance) < .2.40, ps > .121. Similarly, when evaluating models predicting transfer based on explanations in the worked example or transfer problem for each condition separately, only one model was significant. Specifically, for the performance-approach condition, the model predicting transfer based on the number of explanation statements made during the transfer problem itself, was significantly better than a constant-only model, χ^2 (1, N = 35) = 5.07, p = .024. Within this model, the coefficient for the number of explanations was a significant negative predictor of transfer, Wald's χ^2 (1, N = 35) = 4.24, p = .04, *Exp (B)* = .43. This value indicates that, for every explanation statement during this problem that a participant made, the likelihood of transfer decreased by 57%. Specifically, of the 16 participants in this condition who did not generate any explanation statements, 14 transferred correctly (88%). Of the 19 participants who did generate at least one explanation, 10 transferred correctly, (53%).

In summary, the number of self-explanations a participant made throughout the experiment was generally not predictive of transfer, with one exception. The more self-explanation statements that the participants in the performance-approach condition made during the transfer problem, the less likely they were to successfully transfer. This effect may be reflect that self-explaining during the transfer problem is undertaken when a participant is confused or has difficulty understanding the problem. Successfully solving the problem may require already having a firm basis in the procedures needed for the transfer problem, as will be discussed in the next section.

3.4.2 Other Elaborations and Transfer

While explanations were generally not predictive of transfer, it is possible that other types of elaborations may have been. As such, each type of elaborative statements was entered as a predictor into a separate logistic regression, with the likelihood of transfer as a dependent variable. The series of logistic regression analyses found significant effects for a small set of elaborations (see Table 6). In particular, two types of monitoring statements, evaluation (which reflected that a participant was going to double-check their work or their thinking) and open-ended questions, were both marginally significant negative predictors of transfer. Within-problem comparisons were significant negative predictors of transfer, and describing features of the problem were also marginally associated with a lower likelihood of transfer. However, describing problem-solving actions was marginally related to higher likelihood of transfer. In general, it appears that the more monitoring and describing of the problem one did, the lower the likelihood of successfully transferring. Conversely, students who engaged in more descriptions of the problem-solving steps they were taking (or were about to take) were more likely to transfer successfully. It appears that monitoring one's comprehension (i.e., double-checking one's work and asking open-ended questions) or taking basic steps towards comprehending (i.e., comparing aspects of the problem, describing what individual features are) are associated with lower levels of success in transferring, while applying problem-solving steps is associated with higher levels of transfer.

The strongest predictor of transfer was the number of statements students made comparing test material to learning. In particular, when participants referenced the worked example on standardization during the transfer problem, they were much more likely to transfer. In total, 18 students referenced the worked example during the transfer problem; ten of these were in the performance-avoidance condition, six were in the mastery-approach condition, and two were in the performance-approach condition. The likelihood of producing such a statement, for the different conditions was evaluated using logistic regression. This model was significant, χ^2 (2, N = 104) = 6.93, p = .031. The only coefficient which was significantly different from zero was the term describing the difference between the performance-approach and performance-avoidance condition, Wald's χ^2 (1, N = 104) = 5.31, p = .021, *Exp (B)* = .15, indicating that the performance-approach condition was much less likely to produce such statements than the performance-avoidance condition. Of these 18 students who referenced the worked example during the transfer problem, 17 transferred successfully (the one who did not was in the mastery-approach condition).

Table 6

Elaborations as Predictors of Transfer

	Model χ^2	Model *p*-value	Odds Ratio of Coefficient	Coefficient *p*-value
Evaluation	3.67	.055	.81	.058
Question	3.10	.078	.88	.083
Within-Problem Comparison	6.66	.010	.86	.014
Describing Features	2.89	.089	.86	.093
Describing Action	4.18	.041	1.07	.065
Referencing the Worked Example For Transfer	4.83	.028	4.30	.079

NOTE. The "Referencing the Worked Example for Transfer" elaboration was only coded for in the transfer problem, while the other elaborations were coded throughout.

3.5 ACHIEVEMENT GOAL MEASURES

Participant's existing achievement goal orientations for math were assessed as part of the surveys administered at the end of the study. One-way ANOVAs revealed that the experimental conditions responded to these questionnaires differently, even though they asked specifically about their general attitudes. In particular, the ANOVAs revealed differences in how participants responded to the items reflecting performance-approach motivation, F (2, 102) = 7.29, p = .001, and performance-avoidance motivation, F (2, 102) = 10.19, p < .001. Post-hoc analyses indicated that the mastery-approach condition responded significantly lower for both performance-approach and performance-avoidance scales (*M*s = 14.49, 12.66; *SD*s = 4.68, 5.36, respectively) than the performance-approach (*M* = 17.57, 16.09; *SD*s = 3.67, 4.63) or performance-avoidance conditions (*M*s = 17.46, 17.49; *SD*s = 2.95, 3.67). There were no between-group differences for the mastery-approach or mastery-avoidance scales.

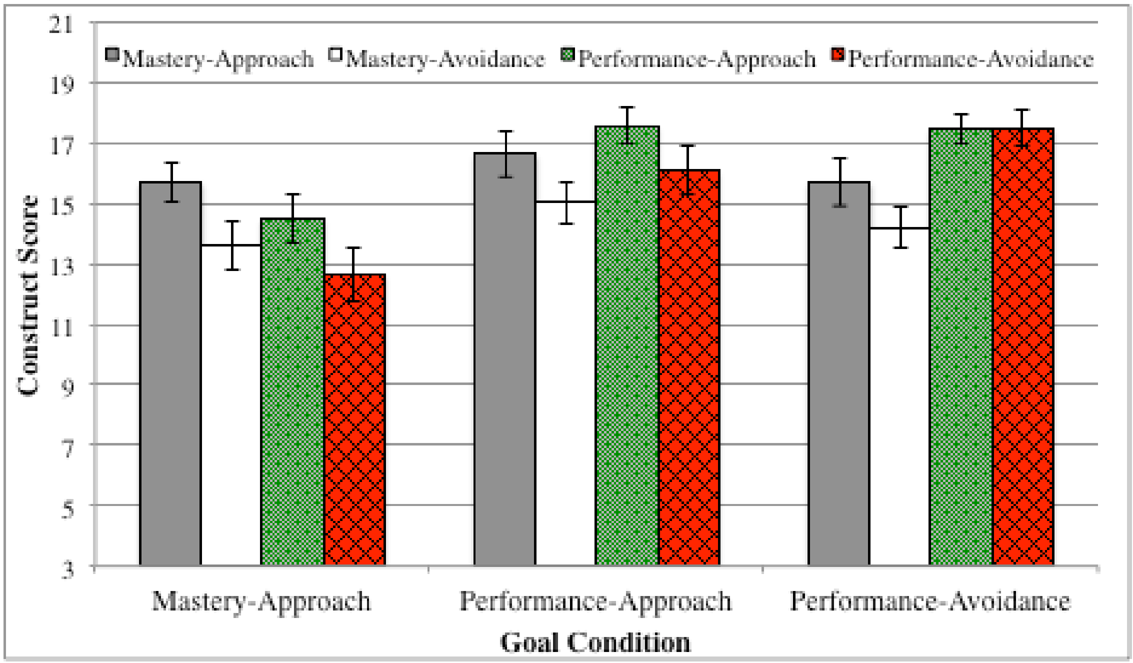

Figure 17. Achievement Goal Questionnaire measures for the three experimental conditions.

4.0 DISCUSSION

The three hypotheses tested in this study examined a behavioral pathway for understanding the benefits of mastery-approach goals on learning. Specifically, it was predicted that mastery-approach goals would produce higher levels of transfer, that these goals would lead to more self-explanation, and that more self-explanation would be predictive of transfer. These hypotheses were not supported. Mastery-approach goals produced the lowest levels of transfer, and did not differ from performance-approach or performance-avoidance in the degree of self-explanations generated, overall. Moreover, the performance-avoidance condition produced the highest level of self-explanation on the worked example, a result that was contrary to expectations. Overall, the number of self-explanations generated was not predictive of transfer for any condition, nor was transfer predicted by self-explanations on any individual component of the study. That being said, the performance-avoidance condition, which did self-explain more on the worked example, went on to have a significantly higher rate of success on the transfer problem than either of the other conditions. In sum, it appears that the condition which showed the most evidence of the hypothesized behavioral pathway of motivation leading to self-explanation leading to better transfer is the condition which received a performance-avoidance goal framing. Potential reasons for this pattern of results will be discussed in this chapter.

Specifically, to understand these results, it is useful to consider the following questions; was the manipulation of goals successful, or did these not actually lead to the goal adoption in the participants that was expected? Why did mastery-approach goals not increase self-explanation? Why did the performance-avoidance condition perform the best on the transfer measure? Are there methodological differences from prior research that could help explain this result? Finally, are there ways to reconcile these results with the extant literature on performance-avoidance goals, which consistently find that performance-avoidance goals are detrimental to academic achievement? These questions will be addressed in order.

4.1 WERE THE MANIPULATIONS SUCCESFUL?

Perhaps the most obvious possible explanation for the results of the current study would be that the manipulations did not work, or that they did not really manipulate mastery-approach goals. The results of this study strongly suggest, however, that the manipulations did indeed influence the goals participants adopted in the experiment, particularly in terms of differentiating mastery from performance goals. Across all of the ways in which the manipulation was checked and goal adoption was measured, the mastery-approach condition produced a consistent pattern of results; this condition reported higher levels of mastery goals and lower levels of performance goals on the manipulation check, the activity questionnaire, the open-ended goal prompt, and the post-test questionnaire. Contrary to expectations, the manipulation even seemed to influence how students responded to the AGQ-R, which is theoretically supposed to measure a different construct (i.e., achievement goals in the context of a math class) than the one manipulated in the study (i.e., achievement goals in the context of the study). The pattern of results indicated that the mastery-approach goal manipulation lowered the degree to which mastery-approach students endorsed performance-approach and performance-avoidance goals, relative to the other conditions.

The pattern was somewhat less clear for differentiating approach and avoidance for performance goals. On one hand, participants were able to cleanly differentiate between these goals on the manipulation check. On the other hand, there was also evidence that manipulations of one goal also increased the adoption of the other. While the manipulation targeted both the criterion component of the goal (i.e., to demonstrate their competence) and the valence (i.e., whether to approach a positive outcome or avoid a negative), the effect on participants' goal adoption indicated that having the criterion led to some amount of goal adoption in both valences. However, it is clear that students in these conditions were adopting a different pattern of goals than students in the mastery-approach condition. Overall, the evidence is strong that the manipulations successfully instilled the appropriate goals.

4.2 MASTERY-APPROACH GOALS AND SELF-EXPLANATION

The current research failed to support the hypothesis that mastery-approach goals would lead to increased levels of self-explanation, compared to the other achievement goals. This is contrary to theoretical expectations of the relations between achievement goals and learning behaviors (e.g., Pugh & Bergin, 2006), as well as to self-reports of student learning behaviors (e.g., Somuncuoglu & Yildirim, 1999), which both point in the direction of a positive relationship between mastery-approach goals and constructive cognitive learning processes. Four possible reasons will be explored that may account for this unexpected result. Specifically, 1) it is possible that mastery-approach goal adoption may not be related to self-explanation at all; 2) that mastery-approach goals are related to the desire to self-explain, but unrelated to the skills necessary to do so; 3) that mastery-approach goals lead to self-explanation when they are paired with other important motivational constructs; 4) or, alternatively, the goal manipulation used in the present work may have produced a different response in participants than the manipulations used in prior experimental manipulations of mastery-approach goals. These will be addressed in the following sections.

4.2.1 Do Mastery-Approach Goals Lead to Self-Explanation?

The first possibility is that mastery-approach goals simply do not lead to self-explanations. While studies have documented a relationship between endorsement of course-based mastery-approach goals and greater self-reported use of strategies like self-explanation, this is no guarantee that the adoption of a task-based mastery-approach goal will lead to self-explanation. That is, while this correlational dependency has been demonstrated, that does not mean there is a causal relationship between the goal and learning behavior. It is possible that the relationship between mastery-approach goals and self-explanation may be due to an unmeasured variable, such as an incremental theory of intelligence (Dweck, 1999) or need for cognition (Cacioppo & Petty, 1982). Alternatively, it may be that self-report measures of learning behaviors are not reliable predictors of actual self-explanation behaviors (see Zhou & Winne, 2012).

4.2.2 Desire vs. Ability to Self-Explain

Another possibility is that mastery-approach goals may lead to a greater *desire* to perform constructive cognitive processing, but that whether or students have the ability to self-explain is orthogonal to this intention. Self-explanation may be a skill that needs to be developed, and, while mastery-approach goals may trigger this behavior in those students capable of doing so, without this skill, the goal will not change learning behavior. The degree to which a person can spontaneously engage in self-explanation may depend on many elements including a sufficient amount of working memory resources, relevant prior knowledge, metacognitive awareness of their understanding, or others. Future research could address this issue by providing explicit guidance and training on self-explanation to students before manipulating the type of goals they adopt for the rest of the experiment. It is also possible that mastery-approach goals do lead to other forms of constructive cognitive processing not explicitly captured by the coding of the protocols done here. This may be because participants were not verbalizing their thoughts completely, even given the instructions and prompting of the experimenter. Additionally, there may be elements of constructive processing that were not coded for that may related to mastery-approach goals, such as schema formation or more complex problem-solving approaches (although this is doubtful, given the lower performance on the transfer item for the mastery-approach condition).

4.2.3 Necessity of Other Motivations

It is also important to consider that mastery-approach goals may, in fact, lead to self-explanation, but only when paired with other important motivational constructs. For example, it is possible that considering the content of the learning material as valuable (as discussed in expectancy-value theories; e.g., Wigfield & Eccles, 2000) while adopting a mastery-approach goal leads to more constructive processing. That is, a student may engage in self-explanation when they both have a goal of understanding and see value in the learning materials they are studying. This would be consistent with Durik and Harackiewicz's (2007) results, which showed that students who have an interest in a topic (which is related to mastery-approach goals; Harackiewicz et al., 2008) respond more favorably to learning materials that are framed as having a high utility value. However, future studies are necessary to untangle what combination of cognitive and motivational factors are necessary for students to engage in spontaneous self-explanation.

4.2.4 Goal Manipulations and Self-Explanation

The present study focused on manipulating goals in a very focused way, targeting the way a participant should define competence for the task, and whether or not they should approach or avoid a given outcome. In the case of mastery-approach goals, this meant guiding participants to focus on trying to develop their skills and understanding of the material presented. While this matches the theoretical definition of a mastery-approach goal, other experimental manipulations of these goals have included other related constructs, and these additional elements may be crucially important for producing self-explanation behaviors. For example, many earlier manipulations included elements to orient participants to a sense of "challenge," which may orient the participants to approach the task in a more constructive manner than an exclusive focus on "improving." The concept of improving may trigger different learning behaviors in participants than the concept of overcoming a challenge, with challenge perhaps leading to a higher degree of self-explanation. This is an open question that future research could address. Additionally, it may be that overcoming a challenge (say, by solving the problems presented correctly) is a more specific goal than improving one's competence; goal specificity has been shown to be an important predictor of behaviors that lead to success (see Locke & Latham, 2002).

4.3 PERFORMANCE-AVOIDANCE GOALS AND TRANSFER

The discussion so far has focused on understanding why mastery-approach goals did not produce the expected pattern of results. Another unexpected outcome of this study is that performance-avoidance goals led to better transfer and to more self-explanation on the worked example. This finding is novel, given the prior literature which generally finds that performance-avoidance goals are only related to lowered performance, achievement and worse affective experiences, and potential reasons for it need to be considered.

Before discussing some potential reasons for this effect, it should be noted that this result is somewhat complicated by the lack of clear differentiation between the performance-approach and performance-avoidance conditions on goal adoption. There were a number of measures that found that students in these conditions had adopted similar goals. Importantly, both the performance-avoidance and performance-approach conditions were higher than the mastery-approach condition in task-based performance-approach goals, and these adopted goals were predictive of transfer for both conditions. So, while it is an open question of why the performance-avoidance condition had the effect of increasing performance-approach goals for the participants, the results may not completely contradict prior research on the role of performance-avoidance goals in learning. That is, performance-avoidance goals themselves may not be related to positive learning outcomes; rather, they may be associated with positive learning outcomes in as much as they also bring about an increase in performance-approach goals. Examined through this lens, the current study provides additional evidence that in some contexts, performance-approach goals can indeed be beneficial to learning and performance. This view is in line with "learning agenda" hypothesis (Senko & Miles, 2005), which predicts that achievement goals may lead students to focus their attention on different aspects of the to-be-learned material. Performance-approach goals, in particular, are predicted to lead students to study in a way that answers, "What do I need to know to do well on an assessment?" In this study, learning the demonstrated procedures well would help them solve the transfer problem correctly with relatively little adaptation necessary. As noted earlier, the performance-avoidance condition was particularly focused on applying the procedure from the worked example to the transfer problem, as evidenced by the larger number of statements comparing the transfer problem to the earlier worked example.

4.3.1 Transfer "Distance" in the Present Study

It is important to consider how and why transfer may not have been as difficult for students in the current study as in prior research with similar materials, as performance goals have been shown to improve performance on more simple tasks, relative to mastery-approach goals. It may be that, to use the terminology of Barnett and Ceci (2002), transfer was "nearer" in this study than in prior research. Specifically, one important difference between the methods of the present study and prior research is how the timing of the experimental protocol was managed. In prior studies, participants had specified amounts of time for each section, and could not proceed until instructed to do so by the experimenter. That is, each part was completed one at a time and if participants finished one of the parts early they had to wait until time was up before they could move on. For example, in the earlier studies the learning phase consisted of the pitching machines problem, instruction delivered via a narrated PowerPoint video, and the track stars problem; participants had exactly 15, 8, and 15 minutes for each, respectively. This was done for the sake of equating everyone on time, as well as to manage running the experiment with multiple participants working individually at the same time.

In the current study, participants had a maximum amount of time allotted for each section, but could proceed through the individual parts at their own pace. For example, in the learning phase of the current study, 20 minutes were allotted for all of the parts, which included the same pitching machines and track stars problems as in prior studies, as well as the worked examples on mean deviation and graphically estimating a standardized score like the tell-and-practice condition of the prior studies. Similarly, they could proceed through the test at their own pace. Also, if a student finished the entire learning phase before the allotted time was finished, the experiment moved on to the test phase, rather than waiting for the full 20 minutes to elapse. One main reason for this change was that it allowed for a more natural talk-aloud protocol to emerge, where participants were always working to get through the packet, and not forcing them to reflect on their prior thinking, as would happen if they were asked to continue thinking aloud about problems they had already completed.

This methodological difference may have had an influence in producing a different pattern of results than the earlier studies. When participants worked on the test phase in earlier studies, they had exactly five minutes for each problem. Because of this, the minimum amount of time between completing the worked example on standardization and the transfer problem was at least ten minutes. In the current study, that time delay could have been (and frequently was) shorter; participants frequently made it through the two intervening problems in less than ten minutes. Temporal continuity is one of the dimensions on which transfer distance can be considered in Barnett and Ceci's (2002) taxonomy, and it was "closer" (i.e., less time elapsed) in the current study than in prior research.

Additionally, this difference may have produced a more holistic sense of the concepts being covered in the entire experiment. When participants worked on each section for a set amount of time, it could create the impression that each is distinct and to be treated separately. Conversely, the structure of the current study may have led participants to treat the entire set of learning materials as tied together, given that they had free rein to move from one to the other as they completed each individual piece. Additionally, they were told that, for the learning section, they could flip back and forth as much as they like, reinforcing the possibility that the material may be linked together for the students. Engle and colleagues (Engle, Lam, Meyer, & Nix, 2012; Engle, Nguyen, & Mendelson, 2011) have shown that contextual cues about the relatively "boundedness" of knowledge can have a large influence on how students process the material, which will influence their ability to transfer.

4.3.2 Potential Benefits of Performance-Avoidance Goals for Learning

Even with the possibility that transfer was nearer in the present research than in prior studies, the counter-intuitive finding that performance-avoidance goals led to the highest levels of transfer must still be accounted for. As such, it is important to also consider reasons that, regardless of transfer distance, performance-avoidance goals may have benefitted student learning in the current study, and, in particular, why this study produced different results than has been demonstrated in prior research.

One possibility is that this study's focus on learning procedures and concepts may have led to a different effect for performance-avoidance goals than prior research has observed, as studies have generally focused on measuring *performance*, rather than learning. Dependent measures from studies in classrooms are generally grades on tests, which may reflect learning but are ultimately an act of performance, and so it may be that the negative relationships observed between performance-avoidance goals and achievement are the results of test-taking or performance issues in particular, such as increased levels of anxiety (Elliot & McGregor, 1999; Skaalvik, 1997).

In fact, one intriguing interpretation of the current study's results is that performance-avoidance goals were adopted by students in that condition, but without any of the negative affect or anxiety that normally accompanies it (e.g., Elliot & Church, 1997). This would be very interesting, as research on the negative effects of "choking under pressure" (e.g., Beilock & Carr, 2005), math anxiety (e.g., Ramirez, Gunderson, Levine, & Beilock, in press), and stereotype threat (e.g., Schmader & Johns, 2003) all suggest that these states use up working memory resources by inserting ruminations on possible negative consequences, worries about poor performance, and self-regulation of negative affect that they are feeling. Stated another way, a performance-avoidance goal tends to increase academic anxiety, and academic anxiety has been shown to be detrimental because it uses up working memory resources. However, while the present study used a successful manipulation of performance-avoidance goals, it did not necessarily succeed in having that goal lead to negative affect and anxiety. That is, the manipulations were constructed in such a way as to ensure that the goal itself was the object of the manipulation, rather than related affective experiences. The manipulation checks that were included in the study all focused on goal adoption, so it is not possible to make strong claims about the affective states of participants in the different goal conditions. While prior experimental studies did not necessarily take into account this difference between performance-avoidance goals and negative affect, there are reasons that they may have succeeded in creating anxiety in their participants, which were not present in the current study. For example, some of these experiments take place in classroom settings (e.g., Elliot et al., 2005), which may be more likely to trigger fears of negative consequences, while others are framed as a way of measuring underlying ability (e.g., Barker, McInerney, & Dowson, 2002), which may cause anxious thoughts about the possibility of having low ability. In the present study, participants believed that performance on these tasks was going to be compared to peers' performance to see if they did worse than others, and that their poor performance would be shown to other students in the future, but participants may not have felt particularly threatened by either of these "negative" outcomes occurring, as they themselves would not necessarily know about it. They may not have felt particularly nervous about this outcome occurring, and so their performance-avoidance goal may have led to energization and goal-directed behaviors, but no anxiety or negative affect. Future research would benefit by paying attention to this distinction; it would help clarify achievement goal theory to pinpoint negative affect and anxiety as the mechanism by which performance-avoidance goals lead to worse performance and achievement.

Given that negative affect or anxiety may not have been induced along with the other aspects of a performance-avoidance goal, the results support a view that performance goals may help learning of less complex material over short periods of time. This would, in some ways, fit with the generally accepted view of how students with performance-focused achievement goals may study and use their knowledge. In particular, the current experiment provided a context where a focus on applying presented formulas directly would lead to success, and this pattern was observed. However, it would not be expected that such learning would be robust, in terms of the duration of learning, nor in how well such learning could be adapted to solve more novel problems. One limitation of the current work was the short duration of the procedure, which precluded measures of retention, but this is something that can be easily addressed in future studies.

Researchers should continue to explore the potential benefits of performance goals on learning and performance of different types of material. Just as research has begun documenting a benefit of mastery-approach goals on knowledge transfer, a research agenda could explore the types of learning outcomes that performance-approach and performance-avoidance goals privilege, as well as examining the mechanisms by which these outcomes occur. While course-based goals have been found to relate to self-reported differences in strategy use (Elliot & McGregor, 1999; Somuncuoglu & Yildirim, 1999), the current study indicates how other methods for observing behaviors (such as talk aloud protocols) may produce different patterns of results. This sort of methodology, as well as other *in situ* measures of goal adoption and learning behaviors, such as those which are integrated into computer-based learning environments (e.g., Bernacki, Nokes-Malach, & Aleven, 2012; Zhou & Winne, 2012), will allow for further refinements and clarifications of achievement goal theory.

5.0 CONCLUSION

The research study presented here continues prior research bridging cognitive psychology of learning with research on student motivation from educational and social psychology (e.g., Belenky & Nokes-Malach, 2012; in press). While prior research had documented a benefit for course-based mastery-approach goals on transfer relative to other course-based achievement goals, the current study finds no benefit for experimentally-manipulated mastery-approach goals. On the contrary, the results showed an unexpected benefit for manipulated performance-avoidance goals relative to mastery-approach goals. Performance-approach goals were also marginally more likely to produce transfer than mastery-approach goals.

These unexpected results raise new questions, and offer many intriguing avenues of research to address them. In particular, this research highlights the need for achievement goal theory to clearly delineate between task-based, proximal achievement goals and course-based goals. While a small number of studies have begun examining the stability of achievement goals over the course of a semester (Fryer & Elliot, 2007; Muis & Edwards, 2009), and even on a task-by-task basis (Bernacki et al., 2012), achievement goal theory generally treats goals that a student has at the beginning of an academic semester and goals that are manipulated by an experimenter as the same. However, it remains open question whether or not these do, in fact, reflect the same construct. Given the results of the current study, which find a benefit for performance-avoidance goals on a learning outcome (a result which few published studies correlating course-based goals to performance have found), and calls from other researchers to examine the use of self-report measures of achievement goals (Richardson, 2004; Zhou & Winne, 2012) it is doubtful that, in the long run, achievement goal theory will be able to conflate these two levels without offering additional clarification.

One possible reason for emphasizing task-based goals in future research is because they may ultimately prove to be more reliable predictors of behavior. It is possible that course-based achievement goals derive their predictive power from their associations with a number of other important constructs, but that they themselves may be, to some degree, epiphenomenal (or, at least, less centrally involved in the processes which lead to achievement). For example, Blackwell, Trzesniewksi and Dweck (2007) found that the degree to which students held incremental theories of intelligence was predictive of both beliefs in the utility of effort and increased mastery-approach goals, but that the effort beliefs were much stronger predictors of the positive learning strategies that directly predicted academic achievement. Similar (unobserved) patterns may be at play in the large body of research that has found links between achievement goals and achievement outcomes.

The present study highlights the need for additional research that unpacks what direct effects achievement goals have on student beliefs and, critically, on actual learning behaviors. The current study illustrated one methodology, talk aloud protocols, for investigating the direct effect of achievement goals on learning behaviors, but others are available (e.g., retrospective think alouds, knowledge tracing via intelligent tutoring systems, etc.), and should be used in future research. In particular, the emerging trend of bridging cognitive science and motivation in empirical research should be encouraged, as both can inform each other in ways that may ultimately improve educational practice. In particular, new theories which integrate insights from each of these fields may have particular power in creating individualized recommendations for a wide variety of learners, based on their unique profiles of prior knowledge, strategy use, and motivation.

Additionally, the current research contributes to the ongoing debate in achievement goal theory about the relative contribution of performance-approach goals to positive learning behaviors and outcomes, and whether or not these goals are to be encouraged. While the results of the present study do not provide guidance on whether or not to “prescribe” performance-approach goal adoption, they do provide additional evidence that performance-approach goals can produce better learning outcomes than mastery-approach goals, in certain contexts. Additionally, given that prior research has documented very few instances of benefits for performance-avoidance goals, the results of the current study are quite surprising. While it may be that, to some degree, the benefit of performance-avoidance goals in the study were due to the degree of performance-approach goal adoption it engendered, further studies using similar manipulations may be able to shed light on precisely why performance-avoidance goals seemed to aid learning and transfer, and whether this effect can be observed in other settings and with other materials. The study reported here was conducted with college students who likely had some familiarity with the covered topics. Perhaps future studies with other populations learning less familiar material will be able to document cleanly whether or not this benefit for performance goals is robust.

While many questions remain for future research to explore, it is clear that the study of motivation in academic learning is a topic of consequence and concern to researchers and educational practitioners alike. Just as advances in the cognitive psychology of memory, learning, and performance have been leveraged to offer concrete guidelines for improving educational practices, theories that bridge our understanding of cognitive processes with aspects of student motivation will be useful for offering concrete recommendations to educators for improved ways of structuring learning activities. The study described here combines theories of student motivation with cognitive psychology methods and measures to explore the complex phenomena of student learning and knowledge transfer, and offers a compelling path for future exploration. By continuing to bridge theoretical and methodological divides, more robust theories of student learning can be proposed, refined, and applied to help educational practice.

BIBLIOGRAPHY

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27,* 669-681.

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26,* 147-179.

Amabile, T.M. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology, 37* (2), 221-233.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84* (3), 261-271.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences, 4,* 167-207.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit and behavioral goals. *Journal of Personality and Social Psychology, 81* (6), 1014-1027.

Barker, K. L., McInerney, D. M., & Dowson, M. (2002). Performance approach, performance avoidance and depth of information processing: A fresh look at relations between students' academic motivation and cognition. *Educational Psychology, 22* (5), 571-589.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128* (4), 612-637.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16* (2), 101-105.

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences, 21* (3), 399-432

Belenky, D. M., & Nokes-Malach, T. J. (in press). Mastery-approach goals and knowledge transfer: An investigation into the effects of task structure and framing instructions. *Learning and Individual Differences.*

Bell, B.S., & Kozlowski, S.W.J. (2002). Goal orientation and ability: Interactive effects on self-efficacy, performance, and knowledge. *Journal of Applied Psychology, 87,* 497-505.

Benware, C.A., & Deci, E.L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21* (4), 755-765.

Bereby-Meyer, Y., & Kaplan, A. (2005). Motivational influences on transfer of problem-solving strategies. *Contemporary Educational Psychology, 30,* 1-22.

Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (April, 2012). *Investigating stability and change in unit-level achievement goals and their effects on math learning with intelligent tutors.* Paper presented at the annual meeting of the American Educational Research Association, Vancouver, BC, Canada.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78* (1), 246-263.

Bong, M. (2004). Academic motivation in self-efficacy, task value, achievement orientations, and attributions. *The Journal of Educational Research 97* (6), 287-298.

Bong, M. (2009). Age-related differences in achievement goal differentiation. *Journal of Educational Psychology, 101* (4), 879-896.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). *How people learn: Brain, mind, experience, and school.* Washington, D.C.: National Academies Press.

Cacioppo, J.T., & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116-131.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Chi, M.T.H., De Leeuw, N., Chiu, M.H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18* (3), 439-477.

Church, M.A., Elliot, A.J., & Gable, S.L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology, 93,* 43-54.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1,* 3, 98-101.

Conroy, D. E., Elliot, A. J., & Hofer, S. M. (2003). A 2 x 2 achievement goals questionnaire for sport: Evidence for factorial invariance, temporal stability, and external validity. *Journal of Sport & Exercise Psychology, 25* (4), 456-476.

Darnon, C., & Butera, F. (2005). Buts d’accomplissement, stratégies d’étude, et motivation intrinsèque: présentation d’un domaine de recherche et validation française de l’échelle d’Elliot et McGregor (2001) [Achievement goals, study strategies, and intrinsic motivation: presentation of a research field and French validation of Elliot and McGregor’s (2001) scale]. *L’Année Psychologique, 105,* 105-131.

Darnon, C., Butera, F., & Harackiewicz, J. M. (2007). Achievement goals in social interactions: Learning with mastery vs. performance goals. *Motivation and Emotion, 31* (1), 61-70.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology, 90* (3), 597-610.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Psychology Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95* (2), 256-273.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34* (3), 169-189.

Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A.J. Elliot & C.S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 52-72). New York, NY: The Guilford Press.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72,* 218-232.

Elliot, A.J. & Fryer, J.W. (2008). The goal construct. In J. Shah & W. Gardner (Eds.) *Handbook of Motivation Science* (pp. 235-250). New York: The Guilford Press.

Elliot, A.J., & Harackiewicz, J.M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A meditational analysis. *Journal of Personality and Social Psychology*, *70* (3), 461-475.

Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *76,* 628-644.

Elliot, A.J., & McGregor, H.A. (2001). A 2 x 2 achievement goal framework. *Journal of the Learning Sciences, 15,* 431-449.

Elliot, A. J., McGregor, H.A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91* (3), 549-563.

Elliot, A.J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100* (3), 613-628.

Elliot, A.J., Shell, M.M, Henry, K.B., & Maier, M.A. (2005). Achievement goals, performance contingencies, and performance attainment: An experimental test. *Journal of Educational Psychology, 97*, 630-640.

Elliott, E.S., & Dweck, C.S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54* (1), 5-12.

Engle, R. A., Nguyen, P. D., & Mendelson, A. (2011). The influence of framing on transfer: Initial evidence from a tutoring experiment. *Instructional Science, 39*, 603-628.

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist, 47* (3), 215-231.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* Boston: MIT Press.

EvoCam (Version 4.0.4) [computer software]. Available from www.Evological.com.

Finney, S.J., Pieper, S.L., & Barron, K.E. (2004). Examining the psychometric properties of the Achievement Goal Questionnaire in a general academic context. *Educational and Psychological Measurement, 64* (2), 365-382.

Ford, K.J., Smith, E.M., Weissbein, D.A., Gully, S.M., & Salas, S. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology, 83* (2), 218-233.

Freud, S. (1920/1990). *Beyond the pleasure principle.* New York: W.W. Norton & Company. (Original work published 1920)

Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology, 99*, (4), 700-714.

Gadgil, S., Nokes-Malach, T.J., & Chi, M.T.H. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction, 22* (1), 47-61.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical comparison. *Journal of Educational Psychology, 95* (2), 393-408.

Gist, M.E., & Stevens, C.K. (1998). Effects of practice conditions and supplemental training method on cognitive learning and interpersonal skill generalization. *Organizational Behavior and Human Decision Processes, 75*, 142-169.

Graesser, A.C., Chipman, P., Haynes, B.C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initative dialogue. *IEEE Transactions in Education, 48*, 612-618.

Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology, 83* (2), 187-194.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85* (3), 541-553.

Harackiewicz, J.M., Barron, K.E., Pintrich, P.R., Elliot, A.J., & Thrash, T.M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology, 94* (3), 638-645.

Harackiewicz, J.M., Durik, A.M., Barron, K.E. Linnenbrink-Garcia, L., & Tauer, J.M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100* (1), 105-122.

Harter, S. (1982). The perceived competence scale for children. *Child Development, 53* (1), 87-97.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52* (12), 1280-1300.

Howell, A. J., & Watson, D. C. (2007). Procrastination: Associations with achievement goal orientation and learning strategies. *Personality and Individual Differences, 43* (1), 167-178.

Hull, C.L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Hulleman, C.S., Schrager, S.M., Bodmann, S.M., & Harackiewicz, J.M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin, 136* (3), 422-449.

Jagacinski, C. M., & Nicholls, J.G. (1984). Conceptions of ability and related affects in task involvement and ego involvement. *Journal of Educational Psychology, 76* (5), 909-919.

Kaplan, A., & Maehr, M.L. (2007). The contribution and prospects of goal orientation theory. *Educational Psychology Review*, 141-187.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*, 523-550.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*, 757-798.

Kozlowski, S. W. J., Gully, S. M., Brown, K. G., Salas, E., Smith, E. M., & Nason, E. R. (2001). Effects of training goals and goal orientations traits on multidimenstional training outcomes and performance adaptability. *Organizational Behavior and Human Decision Processes, 85,* 1-31.

Liem, A.D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology*, *33,* 486-512.

Linnenbrink-Garcia, L., Tyson, D.F., & Patall, E.A. (2008). When are achievement goal orientations beneficial for academic achievement? A closer look at main effects and moderation factors. *International Review of Social Psychology*, *21,* 19-70.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57* (9), 705-717.

Madjar, N., Kaplan, A., & Weinstock, M. (2011). Clarifying mastery-avoidance goals in high school: Distinguishing between intrapersonal and task-mastery standards of competence. *Contemporary Educational Psychology, 36* (4), 268-279.

McClelland, D., Atkinson, J., Clark, R., & Lowell, E. (1953). *The achievement motive*. New York: Appleton-Century Crofts.

Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology*, *34* (4), 265-277.

Murayama, K., Zhou, M., & Nesbit, J. C. (2009). A cross-cultural examination of the psychometric properties of the achievement goal questionnaire. *Educational and Psychological Measurement*, *69,* 266-286.

Murayama, K., & Elliot, A. J. (2009). The joint influence of personal and achievement goals and classroom goal structures on achievement-relevant outcomes. *Journal of Educational Psychology, 101* (2), 432-447.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91,* (3), 328-346.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (in press). Coordinating principles and examples through analogy and explanation. *European Journal of Psychology and Education.*

Ntoumanis, N., & Biddle, S.J.H. (1999). A review of motivational climate in physical activity. *Journal of Sports Sciences, 17* (8), 643-665.

Pajares, F., Britner, S., & Valiante, G. (2000). Writing and science achievement goals of middle school students. *Contemporary Educational Psychology, 25,* 406-422.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 19,* 187-193.

Pintrich, P.R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31* (6), 459-470.

Pintrich, P.R., & DeGroot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance, *Journal of Educational Psychology, 82*, 33-40.

Pugh, K.J., & Bergin, D.A. (2006). Motivational influences on transfer. *Educational Psychologist, 14*, 53-69.

Ramirez, G., Gunderson, E.A., Levine, S.C., & Beilock, S.L. (in press). Math anxiety, working memory and math achievement in early elementary school. *Journal of Cognition and Development.*

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21* (1), 1-29.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Education Psychology, 23,* 90-108.

Richardson, J. T. E. (2004). Methodological issues in questionnaire-based research on student learning in higher education. *Educational Psychology Review*, *16* (4), 347-358.

Richey, J. E., & Nokes-Malach, T. J. (2013). How much is too much? Learning and motivation effects of adding instructional explanations to worked examples. *Learning and Instruction*, *25,* 104-124.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181-210.

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluation metacognitive scaffolding in guided invention activities. *Instructional Science*, *40*, 691-710.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85* (3), 440-452.

Schmidt, A. M. & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology, 56,* 405-429.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*, 129-184.

Senko, C., & Harackiewicz, J.M. (2005). Regulation of achievement goals: The role of competence feedback. *Journal of Educational Psychology, 97*, 320-336.

Senko, C., & Miles, K.M. (2008). Pursuing their own learning agenda: How mastery-oriented students jeopardize their exam performance. *Contemporary Educational Psychology, 33* 561-583.

Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientations, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology, 89,* 71-81.

Smith, M., Duda, J., Allen, J., & Hall. H. (2002). Contemporary measures of approach and avoidance orientations: Similarities and differences. *British Journal of Educational Psychology*, *72,* 155-190.

Somuncuoglu, Y., & Yildirim, A. (1999). Relationship between achievement goal orientations and use of learning strategies. *Journal of Educational Research, 92* (5), 267-277.

Tuominen-Soini, H., Salmela-Aro, K. & Niemivirta, M. (2011). Stability and change in achievement goal orientations: A person-centred approach. *Contemporary Educational Psychology, 36,* 82-100.

Urdan, T. (1997). Achievement goals and the orientation of friends toward school in early adolescence. *Contemporary Educational Psychology, 22*, 165 191

Utman, C.H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychological Review, 1* (2), 170-182.

VandeWalle, D. M., Brown, S. P., Cron, W. L., & Slocum, J. W. (1999).The influence of achievement goals and self-regulation tactics on sales performance: A longitudinal test. *Journal of Applied Psychology, 84,*249 –259.

Van Yperen, N.W. (2003). Task interest and actual performance: The moderating effects of assigned and adopted purpose goals. *Journal of Personality and Social Psychology, 85* (6), 1006-1015.

Van Yperen, N. W., Elliot, A.J., & Anseel, F. (2009). The influence of mastery-avoidance goals on performance improvement. *European Journal of Social Psychology, 39,* 932-943.

Wang, C.K.J., Chatzisarantis, N.L.D., Spray, C.M., Biddle, S.J.H. (2002). Achievement goal profiles in school physical education: Differences in self-determination, sport ability beliefs, and physical activity. *British Journal of Educational Psychology, 72*, 433-445.

Wigfield, A., & Eccles, J.S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.

Wong, R. M. F., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high school mathematics. *Learning and Instruction, 12*, 233-262.

Wolters, C.A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology, 96* (2), 236-250.

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction, 22* (6), 413-419.